Detection of Adulteration in New Zealand High Value Dairy Products using NMR-based Metabolomics: A Chemometrics Approach

Neysa Wendy Setyadi

A thesis submitted for the degree of

Master of Science

at

University of Otago

Dunedin, New Zealand

April 2021

Abstract

New Zealand (NZ) is the world's largest exporter of dairy products. NZ milk products are known for their 'clean and green' reputation with high nutritional properties. In fact, the quality of NZ milk is perceived to be one of the best milk in the world. Cow, goat, and sheep milk are produced in NZ. In fact, NZ goat milk (GM) and sheep milk (SM) are considered as high-value products. The price of goat and sheep milk is 2-3x the price of cow milk (CM) in NZ. These facts put these high value dairy products at risks of fraudulent activities like milk adulteration or counterfeit. Milk can be adulterated with different adulterants including water, vegetable protein, whey, and milk from another species. In other case, expensive milk (e.g., GM, SM) could be adulterated with cheap milk (CM) for the purpose of economic gain. This is also known as economically motivated adulteration (EMA). It is therefore important to develop quick, effective, and robust tools for detection of adulteration. Such techniques must be robust and high throughput,, requires small amount of sample, and highly reproducible. Of particular interest is the application of proton nuclear magnetic resonance (^{1}H-NMR) fingerprinting technique, which meets most of the mentioned requirements.

To date, there is limited information on the metabolite composition of milk produced in NZ. Moreover, NZ milk are not well studied in terms of their compositional properties. Therefore, this thesis is aimed to explore the capability for NMR-based metabolomics technique in detection of adulteration of NZ GM and SM with different concentrations of CM. To achieve this, the study was split into two parts. In the first part, NMR spectroscopy was used to characterise NZ CM, GM, and SM powder to select the discriminant metabolites for each species. In the second part, NMR was used to detect adulteration of GM and SM with different concentrations of CM. Advanced chemometrics (supervised and unsupervised approach) were applied for data interpretation. SPSS and R studio were also employed for statistical analysis.

Overall, NMR fingerprinting technique alongside advanced chemometrics enabled detection of 17, 24, and 23 metabolites present in the water-soluble fractions of CM, GM, and SM, respectively. Out of the identified metabolites, carbohydrates, carboxylic acid, and amino acid were amongst the selected discriminant compounds in CM. In GM, the selected discriminant compounds include amino acid, fatty acid, nucleosides, carbohydrates, and carboxylic acid. Lastly, compounds such as carboxylic acid, carbohydrate, and nucleotide were selected as discriminant markers of SM.

Following characterization, NMR spectroscopy was also successful in identifying potential markers of adulteration in GM and SM with CM. Based on VID feature selection procedure and Tukey's test, phosphocholine was selected as a candidate marker of adulteration of GM with CM. On the other hand, N-acetyl carbohydrates and orotate can be proposed as potential markers of adulteration of SM with CM.

This work is the first study to characterize NZ milk types (CM, GM, SM) using NMR-based metabolomics, and attempt to detect adulteration of NZ GM and SM with different concentration of CM. Overall, NMR-fingerprinting technique was successful in characterising the metabolites present in the different milk type and detecting adulterations. Advanced chemometrics (supervised and unsupervised approach) were also suitable for the interpretation of NMR data, and for identifying discriminants, and in detecting adulterants. Further investigation of different milk fractions (such as lipid) and also the use of other fingerprinting techniques (e.g., LC-Q-TOF-MS, infrared) is needed to support the findings on the present study.

Keywords: NMR, metabolomics, chemometrics, cow milk, goat milk, sheep milk, adulteration detection, metabolites, markers

Acknowledgement

The completion of this project could not have been possible without help and support from numerous individuals. I would extend my sincere gratitude to all of them.

Foremost, I want to thank God without whom nothing is possible.

I would like to extend my deepest gratitude to my primary supervisor, Dr. Biniam Kebede and my co-supervisor, Dr. Dominic Agyei. Their expertise, assistance, guidance, and patience throughout the time of research made it possible for me to complete my master's journey. Thank you so much for your meticulous proofreading.

Besides my supervisor, I would like to thank my mentor Dr. Dan Zhu for sharing her knowledge about the research topic. Thank you for your kind words and suggestions. I wish you best of luck for your future career.

To all the staff and classmates from the Department of Food Science and friends from outside the department; thank you for your word of encouragement and morale support, you guys made my time in University of Otago more fun and enjoyable.

I would like to thank my family; my parents, my sisters for their love, support and prayers that allow me to continue on my postgraduate study. This work would not have been possible without these people.

Table of Contents

Abstract **2**

Acknowledgement **4**

List of Abbreviations **9**

List of Figures **11**

List of Tables **12**

Chapter 1 . Introduction **13**

1.1 Research Strategy: Integrated NMR-based fingerprinting and chemometrics 17

1.2 Thesis Framework 18

Chapter 2 . Literature Review **19**

2.0 Milk: An Overview 19

2.1 New Zealand Dairy Industry 21

2.2 Sources of Milk in New Zealand 22

2.2.1 Cow Milk (CM) 23

2.2.2 Goat Milk (GM) 26

2.2.3 Sheep Milk (SM) 28

2.3 Nutritional Composition of Cow, Goat, and Sheep Milk 30

2.3.1 Milk Protein 31

2.3.1.1 Casein 32

2.3.1.2 Whey Protein 32

2.3.1.3 Non-Protein Nitrogen (NPN) 33

2.3.2 Milk Sugars 34

2.3.2.1 Lactose 35

2.3.2.2 Oligosaccharides 35

2.3.3 Milk Fats 36

2.3.3.1 Milk Fatty Acids Profile 36

2.3.4 Minerals 38

2.3.5 Vitamins 39

2.4 Milk Adulteration 42

2.4.1 Common Milk Adulterants and Public Health Risks 43

2.4.2 Milk Adulteration Incidents 45

2.4.2.1 New Zealand's Involvement in Chinese Milk Scandal 46

2.4.3 Risk of Adulteration of High Value New Zealand Dairy Products 46

2.5 Methods to Detect Milk Adulteration 47

2.5.1 Qualitative Detection of Milk Adulteration 47

2.5.2 Quantitative Detection of Milk Adulteration 49
2.6 The Use Metabolomics in Detection of Milk Adulteration 51
2.6.1 Mass Spectrometry (MS) based milk metabolomics 53
2.6.1.1 LC-MS 54
2.6.1.2 GC-MS 55
2.6.1.3 UPLC-MS 55
2.6.2 Spectroscopy-Based Milk Metabolomics 56
2.6.2.1 Near Infrared (NIR) Spectroscopy 56
2.6.2.2 Fourier Transform Infrared (FT-IR) Spectroscopy 57
2.6.2.3 Nuclear Magnetic Resonance (NMR) based milk metabolomics 57
2.6.3 Summary of Milk-Metabolomics Methods 59
2.7 Chemometrics 63
2.7.1 Unsupervised Approach 64
2.7.1.1 Principal Component Analysis (PCA) 64
2.7.2 Supervised Approach 66
2.7.2.1 Partial Least Squares Regressions Analysis (PLS-R) 66
2.7.2.2 Partial Least Squares Discriminant Analysis (PLS-DA) 67
2.8 Conclusion of Literature Review 68
2.9 Research Gaps 69
Chapter 3 . Objectives, Research strategy and Overall Experimental Approach 70
3.1 Objectives of the study 70
3.2 Research Strategy 70
3.3 Experimental Approach 71
Chapter 4 . Characterisation and Identification of New Zealand Milk (Cow Milk, Goat Milk, and Sheep Milk) 72
4.1 Introduction 72
4.2 Materials and Method for Milk Characterization 74
4.2.1 Samples 74
4.2.2 Reagents 74
4.2.3 Sample Extraction and Preparations for ^{1}H-NMR analysis 74
4.2.4 1H-NMR Experiments 75
4.2.5 1H-NMR Data Pre-processing 75
4.2.6 ^{1}H-NMR Analysis 76
4.3 Multivariate Data Analysis 76
4.3.1 Unsupervised Principal Component Analysis (PCA) 76
4.3.2 Supervised Partial Least Squares-Discriminant Analysis (PLS-DA) 76
4.3.3 Discriminant Markers Selection 77
4.4 Results and Discussion 77

4.4.1 Identification and characterization of NZ cow milk (CM), goat milk (GM), and sheep milk (SM) liquid fraction with ^{1}H-NMR....77
4.4.2 Unsupervised PCA Analysis of NZ CM, GM, and SM....82
4.4.3 Supervised PLS-DA of NZ Cow Milk (CM), Goat Milk (GM), and Sheep Milk (SM)....87
4.4.4 Discriminant Markers Selection with VID and Interpretation....89
4.4.5 ConclusionChapter 4 and Next Steps....98

Chapter 5 . Detection of Adulteration of New Zealand Goat and Sheep Milk with Cow's Milk using NMR Spectroscopy and Chemometrics....100

5.1 Introduction....100

5.2 Materials and Method....101
5.2.1 Samples and Reagents....101

^{1}H-NMR analysis....101
5.2.2 Sample Extraction and Preparations....101
5.2.3 ^{1}H-NMR Experiments....101

5.3 Multivariate Data Analysis....101
5.3.1 Unsupervised Principal Component Analysis (PCA)....101
5.3.2 Supervised Partial Least Square Regression Analysis (PLS-R)....101
5.3.3 Markers Selection with VID....102

5.4 Results and Discussion....103
5.4.1 Adulteration of New Zealand Goat Milk (GM) with Different Concentration of Cow Milk (CM)....103
5.4.1.1 Result from 1H-NMR Spectra of Adulterated GM....103
5.4.1.2 Detecting the Adulteration using Chemometrics with Principal Component Analysis (PCA)....106
5.4.1.3 Detecting the Adulteration using Chemometrics with Partial Least Square Regression Analysis (PLS-R)....106
5.4.1.4 Discriminant Markers Selection of Adulterated Goat Milk (GM)....108
5.4.2 Adulteration of NZ Sheep Milk (SM) with Different Concentration of Cow Milk (CM)....111
The result for the detection of adulteration of NZ SM with CM is explained in the following sections:5.4.2.1 Result from 1H- NMR Spectra of Adulterated SM....111
5.4.2.2 Detecting the Adulteration using Chemometrics with Principal Component Analysis (PCA)....113
5.4.2.3 Detecting the Adulteration using Chemometrics with Partial Least Square Regression Analysis (PLS-R)....113
5.4.2.4 Discriminant Markers Selection of Adulterated Sheep Milk (SM)....115
5.4.3 Summary of Chapter 5....117

Chapter 6 General Discussion, Conclusion, Limitations, and Future Outlook....118

6.1 General Discussion....118

6.2 Conclusion....119

6.3 Limitations of the current study and Future Outlook 120

References 122

Appendix A 155

List of Abbreviations

Milk Types

CM	Cow Milk
GM	Goat Milk
SM	Sheep Milk

Analytical Methods

CE	Capillary Electrophoresis
ELISA	Enzyme Linked Immunosorbent Assay
FT-IR	Fourier Transform Infrared Spectroscopy
GC-MS	Gas Chromatography - Mass Spectrometry
HPLC	High Performance Liquid Chromatography
LC-MS	Liquid Chromatography - Mass Spectrometry
MS	Mass Spectrometry
NIR	Near-infrared spectroscopy
NMR	Nuclear Magnetic Resonance
PCR	Polymerase Chain Reaction
UPLC	Ultra-Performance Liquid Chromatography

Statistical Methods

LV	Latent Variable
PC	Principal Components
PCA	Principal Component Analysis
PLS	Partial Least Square
PLSDA	Partial Least Square - Discriminant Analysis
PLSR	Partial Least Square - Regression
VID	Variable Identification

Milk Compositions

CLA	Conjugated Linoleic Acid
DHA	Docosahexaenoic Acid
EPA	Eicosapentaenoic Acid
FA	Fatty Acids
FAA	Free Amino Acids
MUFA	Monounsaturated Fatty Acids
NPN	Non-Protein Nitrogen
PUFA	Polyunsaturated Fatty Acids
SCC	Somatic cells count
SFA	Saturated Fatty Acids

Reagents

ACN	Acetonitrile
D_2O	Deuterium Oxide
PB	Phosphate Buffer
TSP	Sodium 3-(trimethylsilyl)(2,2,3,3,H4)propionate

Miscellaneous

CMPA	Cow Milk Protein Allergy
CVD	Cardiovascular Diseases
EMA	Economically Motivated Adulteration
GHG	Green House Gas
NZ	New Zealand

List of Figures

Figure 1.1. Article records on NMR-based metabolomics and NMR-milk metabolomics publications using the keyword NMR and metabolomics or NMR and milk metabolomics[1,2] 16
Figure 1.2. Research Strategy to integrate NMR fingerprinting, advanced chemometrics, and biomarker selections 17
Figure 1.3. Framework of the present thesis 18
Figure 2.1. Top 10 export destination of NZ Dairy Product (As adapted from MPI (2019)) 22
Figure 2.2. New Zealand Cow Breeds during the year of 2019 24
Figure 2.3. Advantages and Disadvantages of Different Metabolomics Method (targeted, pseudo-targeted, and untargeted methods) 53
Figure 3.1. Experimental approach for characterization and detection of adulterations of New Zealand cow, goat, and sheep milk 71
Figure 4.1. **^{1}H**-NMR Spectra of NZ CM, GM, and SM (removed lipid fractions) 78
Figure 4.2. Venn Diagram for metabolites compounds found in CM, GM and SM 82
Figure 4.3. Cumulative Variance Graph - NZ CM, GM, and SM Characterization (PCA) 83
Figure 4.4. Root Mean Square Error of Cross Validation (RMSECV) Graph – NZ CM, GM, and SM Characterization (PCA) 83
Figure 4.5. A PCA Score Plot on the Characterization of NZ CM, GM, and SM 84
Figure 4.6. PCA loading plots for CM, GM, and SM metabolites 85
Figure 4.7. A PCA biplot illustrating the compound variance between NZ CM, GM and SM 86
Figure 4.8. Cumulative Variance Graph – Characterization of NZ CM, GM and SM (PLS-DA) 87
Figure 4.9. Root Mean Square Error of Cross Validation Graph - Characterization of NZ CM, GM, and SM (PLS-DA) 87
Figure 4.10. PLS-DA scores plot on the characterization of NZ CM, GM, and SM 87
Figure 4.11. PLS-DA loading Plot for the characterization of NZ CM, GM, and SM 88
Figure 4.12. PLS-DA bi-plot on characterization of NZ CM, GM and SM 89
Figure 4.13. Positive discriminant compounds selected through VID procedures for cow milk (CM) 91
Figure 4.14. Positive discriminant compounds selected through VID procedures for goat milk (GM) 93
Figure 4.15. Positive discriminant compounds selected through VID procedures for sheep milk (SM) 97
Figure 5.1. NMR Spectra of NZ GM adulterated with different concentration of CM (1%, 2%, 4%, and 8%) 104
Figure 5.2. Cumulative Variance Graph - GM adulterated with CM (1%, 2%, 4%, and 8%) (PLSR) 106
Figure 5.3. Root Mean Square Error of Cross Validation Graph - GM adulterated with CM (1%, 2%, 4%, and 8%) (PLS-R) 106
Figure 5.4. PLS-R bi-plot showing the change in the NMR metabolite profile of GM due to the adulteration with different concentration of CM (1%, 2%, 4%, and 8%) 107
Figure 5.5. Discriminant compounds for comparison of unadulterated GM and adulterated GM with different concentration of CM (1%, 2%, 4%, and 8%), selected through VID procedure[1] 109
Figure 5.6. NMR Spectra of unadulterated SM with SM adulterated with different concentration of CM (1%, 2%, 4%, and 8%) 111
Figure 5.7. Cumulative Variance Graph - SM adulterated with CM (1%, 2%, 4%, and 8%) (PLSR) 113
Figure 5.8. Root Mean Square Error of Cross Validation Graph - SM adulterated with CM (1%, 2%, 4%, and 8%) (PLS-R) 113

Figure 5.9. PLS-R Bi-plot for detecting SM adulteration with different concentration of CM (1%, 2%, 4%, and 8%) 114
Figure 5.10. Discriminant compounds for comparison of unadulterated SM and adulterated SM with different concentration of CM (1%, 2%, 4%, and 8%), selected through VID procedure[1] 116
Figure A.1. PLS-R score plot for adulterated GM adulteration with different percentage of CM (1%,2%,4%, and 8%) 155
Figure A.2. PLS-R loadings plot for adulterated GM adulteration with different percentage of CM (1%,2%,4%, and 8%) 155
Figure A.3. PLS-R score plot for detecting SM adulteration with different percentage of CM (1%,2%,4%, and 8%) 156
Figure A.4. PLS-R loadings plot for detecting SM adulteration with different percentage of CM (1%,2%,4%, and 8%) 156

List of Tables

Table 2.1. Gross composition of cow, goat and sheep milk[a] 31
Table 2.2. Casein fractions of cow, goat, and sheep milk[a,b] 32
Table 2.3. Whey protein constituents in cow, goat, and sheep milk[a-b] 33
Table 2.4. Free amino acids composition (mg/kg) in the proteins of cow, goat, and sheep milk[a,b] 34
Table 2.5. Fatty Acids Profile (g/ 100g) in cow, goat, and sheep milk[a,b] 38
Table 2.6. Minerals concentration in cow, goat, and sheep milk[a,b] 39
Table 2.7. Dietary reference intake for vitamins, its main role, and consequences of deficiencies in human health[a] 41
Table 2.8. Vitamin content of cow, goat, and sheep milks[a,b] 42
Table 2.9. Common milk adulterants, uses and their public health risks 44
Table 2.10. Rapid Qualitative Detection of Different Adulterants in Milk 47
Table 2.11. Metabolomics terms and definitions 51
Table 2.12. Comparison of NMR and MS-based Methods for Metabolomics Analysis 60
Table 2.13. Summary of Milk-based Metabolomics Studies Reported in The Literature[1,2] 61
Table 2.14. Chemometrics terms and definition (As modified from Ellis et al. (2012)) 63
Table 2.15. Chemometrics modelling methods and their uses 64
Table 4.1. Metabolites assignment from ^{1}H-NMR spectra of water-soluble fraction from NZ CM, GM, and SM 81
Table 4.2. Discriminant marker compounds selected for cow milk (CM)[1] 90
Table 4.3. Discriminant marker compounds selected for goat milk (GM)[1] 93
Table 4.4. Discriminant marker compounds selected for sheep milk (SM)[1] 96
Table 5.1. Metabolites assignment from ^{1}H-NMR spectra of NZ GM adulterated with different concentration of CM (1%, 2%, 4%, and 8%) 105
Table 5.2. Potential markers for detecting the adulteration of GM with CM selected by VID procedure[1] 108
Table 5.3. Metabolites assignment from 1H-NMR spectra of NZ SM adulterated with different concentration of CM (1%, 2%, 4%, and 8%) 112
Table 5.4. Potential markers for detecting the adulteration of SM with CM selected by VID procedure[1] 115

Chapter 1 . Introduction

Milk is a nutrient-dense food, consumed by a majority of the world's population. In its natural form, milk is an excellent source of proteins, carbohydrates, fats, vitamins, and minerals (Poonia et al., 2016). According to Neumann, Harris, and Rogers (2002) proteins in milk are considered as the highest quality, as they contain a full amount of essential amino acids followed by their resemblance to the proteins present in the human body. Milk also contains a wide range of micronutrients (especially zinc, calcium, potassium, phosphorous, and magnesium) existing in an easily absorbable form, thus, making it nature's most complete food (Górska-Warsewicz, Rejman, Laskowski, & Czeczotko, 2019).

Milk consumption is associated with improved bone health and bone density, and it also reduces the risk of bone fractures in children (Caroli, Poli, Ricotta, Banfi, & Cocchi, 2011; Park, 2009). Depending on the age of the consumer, milk can contribute around 20-28% of dietary reference intake (DRI) of protein, and around 52-65% of the requirement for calcium (Rozenberg et al., 2016).

Because of the perishable nature of milk, most dairy products in the world are consumed in the country or region in which they are produced (Shadbolt & Apparao, 2016). In New Zealand (NZ) however, only 5 % of the produced milk is consumed within the country; whilst 95% of its products are exported to more than 100 countries around the world (TDB, 2018). This makes NZ the world's largest dairy exporter and the 8th-largest dairy producer in the world (DCANZ, 2020). Additionally, not all milk in NZ is coming from cows; milk from goats, sheep, and deer is also produced in NZ. Out of four milk types produced in NZ, cow milk (CM) is the cheapest and most popular. Recently, there has been an increase in the demand for NZ goat milk (GM) and sheep milk (SM) globally, while the awareness of deer milk (DM) is relatively limited (Miller & Lu, 2019).

In general, NZ milk products are known for their clean and green reputation with high nutritional properties (MFE, 2001). In fact, the quality of NZ milk is perceived to be one the highest in the world (NZMP, 2018). Furthermore, NZ milk is said to be richer in omega-3 fats, β-carotene, and conjugated linoleic acid (CLA) (Rugoho, Liu, & Dewhurst, 2014). For this reason, NZ milk is quite expensive compared with the milk from another country. It costs around $6.75 per kg of milk solids for CM, $17 per kg of milk solids for GM, and $16 per kg of solids for SM (Hall, 2019). Thus, the high quality and high price of NZ milk provides a motivation for milk adulteration.

As the second-most adulterated food product in the world, after olive oil; there are many ways that milk gets adulterated (Moore, Spink, & Lipp, 2012). Substances that can be added into milk includes water, salt, sugar, detergents, starch, melamine, urea, and vegetable oils (Nascimento, Santos, Pereira-Filho, & Rocha, 2017). Whilst milk adulteration is commonly intended for economic gain, milk adulterants are usually associated with negative health effects, which, in some cases, could result in death of consumers (Xiu & Klein, 2010). Furthermore, milk adulteration will destroy the reputation of the dairy brand, causing a significant economic loss to the company (Everstine, Spink, & Kennedy, 2013). Thus, in order to maintain NZ's dairy brand reputation, it is important to have an effective and robust analytical method to detect milk adulteration.

There are many ways to detect adulterants in milk. One of these is the qualitative detection method that focuses on the colour changes via chemical reactions. The qualitative detection method is easy and straightforward to apply. However, it is only valid for limited types of compounds with limited ranges of concentration (Azad & Ahmed, 2016). Other methods such as high-performance liquid chromatography (HPLC) (Cheah & Fang, 2020), sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (Yang, Zheng, Soyeurt, Yang, & Wang, 2019), capillary electrophoresis (CE) (Trimboli et al., 2019), and ion mass spectrometry could also be used to detect adulterants in milk. Even so, these methods are highly complex and time-consuming. Another method such as gas chromatography-flame ionization detector (GC-FID) is also good in adulteration detection (Pizzo et al., 2018). Nevertheless, as with the other traditional methods, it is destructive and results in irreversible changes in the tested samples.

For the reasons stated above, it is important to employ other technique that is rapid and capable of detecting a wide range of metabolites, even at a minute concentration, such as metabolomics. Metabolomics is an emerging field that focuses on the study of low molecular weight compounds in a biological sample (Wishart, 2008). In food analysis, metabolomics can be used to investigate how a product changes with food processing, for detection of food contaminants and adulterants, and for understanding the correlation between flavour compounds and product liking (Mermelstein, 2019).

Based on the technique and its detection method, metabolomics can be classified into three types: untargeted, pseudo-targeted, and targeted (Roberts, Souza, Gerszten, & Clish, 2012). Untargeted metabolomics is the comprehensive analysis that focuses on the detection of as

many metabolites as possible without having to quantify them (Gertsman & Barshop, 2018). Pseudo-targeted metabolomics is based on the combination of the advantages of untargeted and targeted metabolomics, where it focuses on the quantification of unknown compounds and partly known compounds (Xu et al., 2019). Targeted metabolomics focuses on the detection and quantification of known metabolites. Since the purpose of thesis was to protect high-value NZ dairy products from adulteration, it was important to employ the method that can detect as many metabolites as possible. As a result, the untargeted metabolomics method was selected as the main approach in this study.

To date, NMR spectroscopy along with LC-MS and GC-MS are the three most commonly used analytical methods in untargeted metabolomics (Sundekilde, Larsen, & Bertram, 2013). Whilst the application of LC-MS and GC-MS methods are becoming increasingly popular, there is still considerable interest in the use of NMR spectroscopy as a tool for metabolomics studies (Emwas et al., 2019).

There are currently a total of 907 publications on NMR-metabolomics since the year 2000, while there is only a total of 21 publications on NMR-based milk metabolomics (**Figure 1.1**). This means, the application of NMR-based metabolomics in milk studies is still relatively new and the application of NMR in dairy research had only become popular in 2013. **Figure 1.1** illustrates the steadily increasing number of applications of NMR in metabolomics studies and its application in milk-based studies.

NMR is a non-biased method that is capable of detecting and characterizing compounds that are less tractable in LC-MS analysis (Takis, Ghini, Tenori, Turano, & Luchinat, 2019). NMR is also highly reproducible, making it suitable for for a large –scale metabolomics studies (Guennec, Giraudeau, & Caldarelli, 2014).

Sundekilde, Poulsen, Larsen, and Bertram (2013) used NMR to investigate the association between somatic cell count (SCC) in milk with milk metabolites. In another study, NMR was used to obverse changes of metabolites during lactation in human milk (Wu et al., 2016). NMR was also used as a tool to identify milk authenticity based on metabolites analysis (Li, Yu, et al., 2017) More recently, Cui et al. (2019) used NMR to identify biomarkers in reconstituted milk. Various NMR studies were found, but not on the detection of milk adulteration. Other than that, there is little or no work on the characterisation of NZ milks (cow, goat, sheep) with NMR spectroscopy. Thus, it will be interesting to explore the potential of NMR-based

metabolomics in the detection of adulteration in NZ milks. For this reason, NMR-based metabolomics technique was chosen as the primary technique in the present study.

In the present study, metabolite fingerprinting was performed to analyse and identify the presence of metabolites in different types of milk and to identify the metabolites compound. Since NMR data is rich in information, it is important to apply advanced chemometrics for data exploration and interpretation.

Overall, the aim of the thesis is to develop a technique for quick detection of adulteration. Specifically, to find out whether NMR-based metabolomics methods combined with advanced chemometrics is capable of detecting low concentrations of adulterants in NZ GM and SM. Details regarding on the specific objectives of the thesis is also mentioned on Chapter 3.

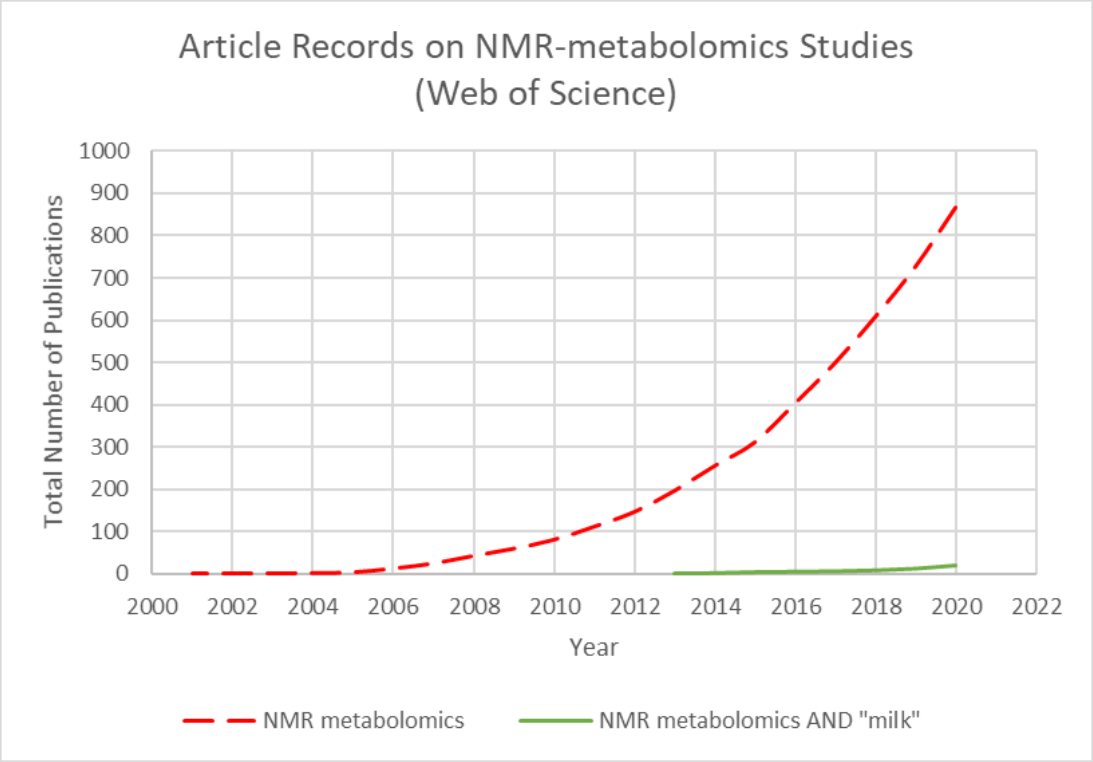

Figure 1.1. Article records on NMR-based metabolomics and NMR-milk metabolomics publications using the keyword NMR and metabolomics or NMR and milk metabolomics[1,2]

[1]Sourced from the Web of Science (http://apps.webofknowledge.com/)
[2]Keywords used in the study were "NMR metabolomics" AND "milk" and "NMR metabolomics"

1.1 Research Strategy: Integrated NMR-based fingerprinting and chemometrics

To achieve the aim of the current study, a research strategy was created (**Figure 1.2**).

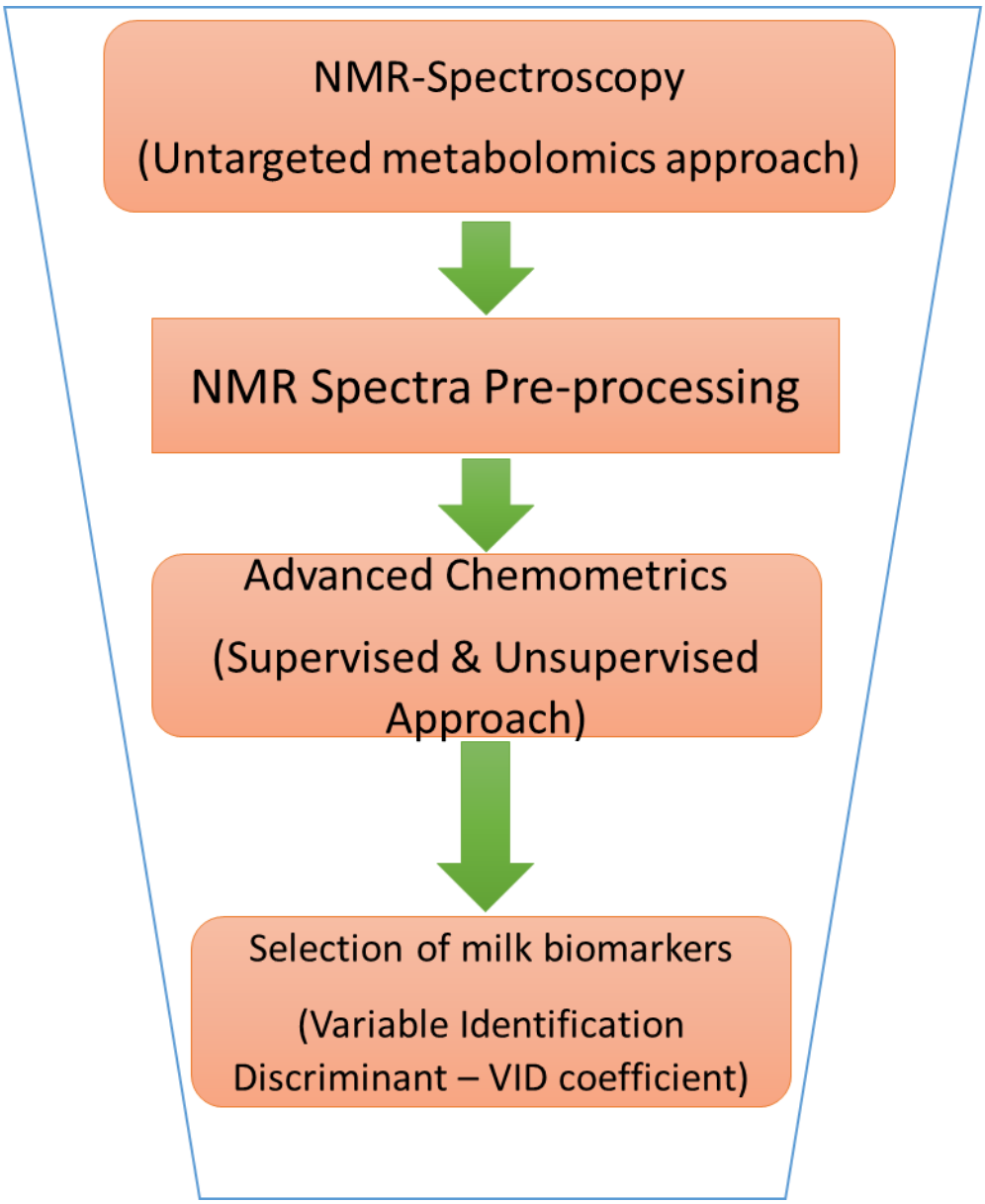

Figure 1.2. Research Strategy to integrate NMR fingerprinting, advanced chemometrics, and biomarker selections

1.2 Thesis Framework

The overall framework of the thesis can be seen on **Figure 1.3**. The first chapter, **Chapter 1** covers the general introduction of the topic, the aim of the thesis, research strategy, and framework. **Chapter 2** covers the literature study focusing on the nutritional information of milk, risks of adulterations in milk, analytical techniques to characterize and detect adulteration in milk, and finally data analysis technique. Based on the literature review and research gaps, the specific objective and the experimental approach was proposed in **Chapter 3.** To achieve the general objective, the experimental sections in this study are divided into two parts. The first part about characterization and identification of different NZ milk species with NMR-based metabolomics is covered in **Chapter 4.** The second part, which focuses on the detection of adulteration of NZ GM and SM with different concentration of CM is covered in **Chapter 5.** Finally, the general conclusions, the implication and recommendations of the study are summarised and presented in **Chapter 6**.

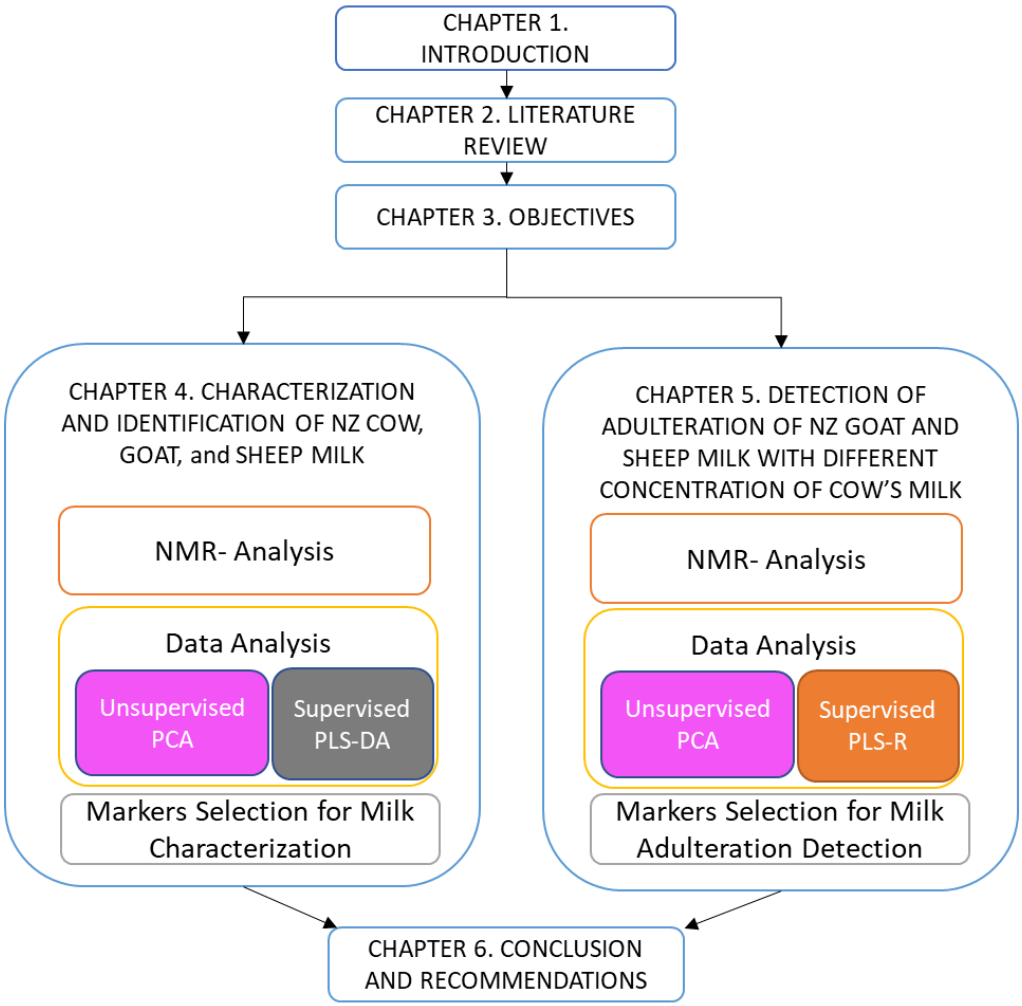

Figure 1.3. Framework of the present thesis

Chapter 2 . Literature Review

The aim of this MSc project is to develop NMR-based metabolomics approach to detect adulteration in high-value NZ dairy products (goat and sheep milk). Therefore, it was deemed important to conduct a literature review to gain understanding of the milk composition and the methods for detecting adulteration.

Accordingly, the literature study covers topics such as an overview of global milk production, the NZ dairy industry and its milk sources, milk composition, milk adulteration, metabolomics approach for detection of milk adulteration, and statistical methods for analysing metabolomics data set.

2.0 Milk: An Overview

The U.S. Food Drugs and Administration (FDA) defined milk as a "liquid food produced from the lacteal secretion obtained from one or more healthy milch animals such as cows, water buffalo, camels, goat, sheep, and others". The range of milk products coming from different countries and even within the same countries sometimes varies significantly. This is mainly influenced by the social and cultural conditions of the country, available dairy processing facilities, and the market demand (OECD & FAO, 2019a). For instance, camel milk is predominantly consumed in Middle Eastern countries and it is not available in NZ.

Milk is one of the most important nutrient sources for the human diet, containing substantial amounts of macro-and micro-nutrients essential for human growth (Wiley, 2007). As a versatile product, milk can be processed and transformed into other essential ingredients used to manufacture food products. Such products derived from milk include butter, cream, cheese, yoghurts, ice cream, and probiotics. Milk and milk products are nutrient-dense foods. They can provide nine essential vitamin and minerals, including, vitamin A, D, B12, potassium, phosphorous, protein, calcium, niacin, and riboflavin. They perform various health roles in humans (Iqbal, 2017).

Milk and milk products contain several nutrients required to build strong bones, especially during childhood. Several studies have shown that the consumption of dairy products reduces a person's risk of developing osteoporosis, helps prevent cardiovascular disease in elderly people, and reduces risk of obesity and type 2 diabetes (Gijsbers et al., 2016; Thorning et al., 2016). Milk consumption is also positively associated with foetal growth and infant birthweight in pregnant women (Kalkwarf, Khoury, & Lanphear, 2003; Thorning et al., 2016; Vogel et al., 2017). For these reasons, the consumption of milk is encouraged within the dietary guidelines

of many countries, as means to attain sufficient nutritional status. The U.S. Department of Agriculture (USDA) and U.S. Department of Health and Human Services (HHS) have suggested a recommended intake of 2 cups equivalent of milk per day for children ages 2–3 years, 2 ½ cups equivalent per day for children ages 4–8 years, and 3 cups equivalent for adolescents ages 9–18 years and adults.

In 2019, the global dairy market in the world had reached more than 1 trillion NZD and by 2030 the milk demand in the world is expected to increase by 35% (FAO, 2020a). Milk is one of the most produced and valuable commodities worldwide, where it is consumed and produced in every country (Ritchie & Roser, 2017). It ranks amongst the top five agricultural commodities in both quality and value terms. Around the world, milk from cows represent 82.7% of the global milk production, followed by buffaloes' milk with 13.3%, goat's 2.3%, sheep 1.3%, camel's 0.4%, and other sources < 0.1% (FAOSTAT, 2014).

More than 6 billion people consume milk and milk products and over 200 billion litres of milk are consumed every year (Sugrue, Tobin, Ross, Stanton, & Hill, 2019). By volume, liquid milk is the most consumed dairy product in both developing and developed countries (G. M. Singh et al., 2015). With the increase in world population and income per capita, the world milk industry is growing very fast. There has always been an increase of demands for better quality of food from animal sources and dairy products (Flachowsky, Meyer, & Südekum, 2017). The world milk production is estimated to increase by 177 million tonnes by 2025, with an average growth rate of 1.8% per annum. Additionally, consumption of dairy products per capita is also projected to increase between 0.8 – 1.7% per year in developing countries, and 0.5 – 1.1% in developed countries (OECD & FAO, 2016). According to OECD and FAO (2019b), milk provides 3–4% of dietary energy supply in Asia and Africa and around 9% in North America, Europe, and Oceania. Additionally, it is estimated that milk contributes to 19% of dietary protein supply in Europe, America, and Oceania, and 8% in Asia and Africa. This proves that milk is widely consumed around the world.

In 2019, the top-ten milk producers by countries were United States of America, India, China, Brazil, Germany, Russia, France, NZ, Turkey, and United Kingdom (FAO, 2020a; Loeschen, 2019). Among these countries, NZ had the highest milk surpluses, followed by United States of America, Germany, and France. NZ exports nearly 95% of its dairy products to more than 100 countries, accounting for approximately 35% of the world trade in dairy products

(DCANZ, 2020; NZIER, 2017). With this, NZ plays a significant role in the global dairy industry.

2.1 New Zealand Dairy Industry

The dairy industry is defined as an enterprise that deals with the harvesting and processing of animal milk for human consumption (Faye & Konuspayeva, 2012). In NZ, the dairy industry is one of the major sectors in the country's economy in which it contributes around 3.5% ($7.8 billion) of the country's total GDP. The dairy farming plays a crucial role in the country's economy where it contributes 28% of the total goods exports (2017–2018), greater than the meat, wine, and wood sector (Ballingall & Pambudi, 2017). According to DairyNZ (2019), NZ's dairy industry reportedly feeds more than 100 million people worldwide and there are about 46,000 employees in the dairy industry in NZ. From these numbers, 34,000 are on-farms sites, while 12,000 are involved in processing and wholesale of dairy products.

In the world, NZ has the highest level of dairy-sufficiency. This is supported by the natural conditions of the country and its small population. According to Ledgard, Judge, Smeaton, and Boyes (2010), a more efficient dairy system means the product would have lower carbon footprint. This implies the average dairy product from NZ produces less methane compared with the dairy product from other countries.

Generally, most dairy products around the world are consumed in the region or the country where they are being produced, because of the perishable nature of the milk (Shadbolt & Apparao, 2016). In NZ however, as mentioned previously, 95% of dairy products are exported each year. The remainder, 5%, were said to have met the country's domestic requirements for liquid milk products that includes fresh cream, skim milk, full-cream milk, and trim milk (TDB, 2018). This makes, NZ the world's largest dairy exporter and the eight-largest dairy producer in the world (DCANZ, 2020).

Additionally, the source of milk in NZ is not limited to cows. There are several dairy farms focusing on the production of goat, sheep, and deer milk. Even so, the existence of deer milk in the market is still scarce as compared with cow, goat, and sheep milk.

About half of the dairy exports from NZ are value-added products, with the aid of different processing and packaging conditions (Doole, 2014). The reputation of NZ being the world's largest dairy exporter, the reputation of the product manufacturers, and the fact that NZ uses a sustainable farming system contribute to the high value and positive consumer perception of NZ dairy products.

Over the past decade, the dairy export value in NZ has risen significantly owing to the increasing milk solid production and the extensive export destinations. To date, NZ dairy products are marketed to more than 100 countries around the world. As of 2019, the top-10 export destinations for NZ dairy products were China, Australia, Japan, Malaysia, United States, Philippines, UAE, Indonesia, Thailand, and Hong Kong (**Figure 2.1**). NZ exports whole milk and skim milk powder mostly to Asia, cheese to Asia and Australia, butter to Europe, and other products such as casein powder to United States (OEC, 2017).

Figure 2.1. Top 10 export destination of NZ Dairy Product (As adapted from MPI (2019))

2.2 Sources of Milk in New Zealand

Milk and dairy products are considered essential to achieve a balanced diet, possessing a wide range of bioactive components (Hsieh et al., 2015). Milk is consumed by infants, children, adults, and even the elderly worldwide as a high-quality source of protein and calcium (Yadav et al., 2020). The source of milk itself could come from three different types: human, animal, and plants. Human milk is more suitablefor human consumption compared to milk from animal and plants. However, human females only produce milk during the lactation stage and human breast milk is not distributed commercially. Breast milk is only consumed by infants for around six months to a year, depending on the mother and the conditions of the infant (da Costa et al., 2010). Thus, as an alternative, human beings could consume either animal or plants milk as these are both distributed and produced commercially.

The animal milk suitable for human consumption can be sourced from cow, buffalo, yak, camel, horse, donkey, sheep, goat, and deer. On the other hand, the sources for plant milk include almond milk, rice milk, oats milk, flax milk, hemp milk, coconut milk, and soy milk (Sethi, Tyagi, & Anurag, 2016). Although there are many sources of milk, in terms of nutritional value and taste, animal milk is more superior compared with the plant-based milk (Haas, Schnepps, Pichler, & Meixner, 2019). For example, Palacios, Badran, Drake, Reisner, and Moskowitz (2009) reported higher overall liking in lactose-free CM as compared to soy-based milk, regardless of if the consumer were lactose intolerant. Another study in Switzerland had analysed 45 plant-based milk products obtained from supermarkets in connection with their nutrient content. This study reported that the substitution of cow milk with plant-based milk had resulted in reduced intake of calcium, proteins, minerals, and several types of vitamins (Kopf-Bolanz & Sousa, 2017).

Historically, humans are the only animals known to consume milk coming from another species. It is deemed a unique behaviour that arose during Neolithic Revolution 12,000 years ago (Knopfler, 2016). Since then, milk coming from livestock animals has been a significant source of nutrients for humans. Currently, cow milk is the most commonly consumed and produced milk in the world, dominating the global milk production owing to the large cattle population around the world (S. D. Kalyankar, Khedkar, Patil, & Deosarkar, 2016). Even so, people from other parts of the world might consume milk coming from other animal sources, depending on the availability and their nutritional benefits. For example, in NZ, the three popular animals used as milk sources are cow, sheep, and goat.

Since the focus of this study is on NZ milk sources, the details about each source and their benefits are discussed more in the following sections: **Section 2.2.1 Cow Milk**, **Section 2.2.2 Goat Milk**, and **Section 2.2.3 Sheep Milk**.

2.2.1 Cow Milk (CM)

NZ is a country with a population of approximately 5 million people and almost 6.3 million milking cows (11,400 herds), meaning that there are more dairy cows than people (Sneddon et al., 2015; StatsNZ, 2021). The national herd in New Zealand is made up of three main breeds: Holstein-Friesian (33%), Jersey (8.6%) and their crossbred (48.5%). The rest are minority breeds, including Ayrshire (0.5%), and others (9.3%) (Milking Shorthorn, Guernsey, Brown Swiss, and their crossbred) (DairyNZ, 2019; Sneddon et al., 2015).

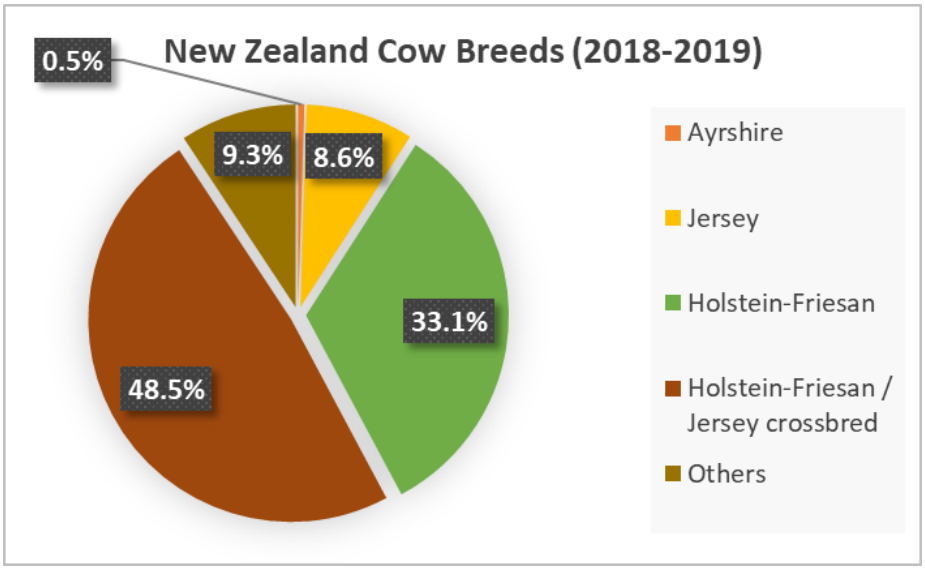

Figure 2.2. New Zealand Cow Breeds during the year of 2019

As the most studied milk in the world, CM has always been part of the human diet for thousands of years (Haas et al., 2019). Humans started to consume CM as early as their infancy stage, usually around the end of breastfeeding (Leung & Sauve, 2003). Generally, CM is chosen for the substitution of breast milk because of its availability and its high content of calcium, protein, and vitamins (A, B2, B12) (Haas et al., 2019). As a result of its high nutrition and health benefits, CM is included in many dietary recommendations regardless of age group.

Some studies have shown that CM consumption improves bone health, density, and also reduces the risk of having bone fractures in children (Caroli et al., 2011). People who consumes CM were also said to have reduced risk of cardiovascular diseases and type 2 diabetes compared with those who don't consume cow milk at all (Delvina, 2017; Gijsbers et al., 2016). Additionally, people who consumed milk and milk products regularly in their adult years have lower risk of having osteoporosis in postmenopausal women and elderly (S.-J. Park, Jung, Kim, & Lee, 2018; Rangel et al., 2016).

Other than health benefits, CM is associated with good taste and high taste preference. Harry (1977) did a study on taste preference in milk on human ontogeny, where he found that new-borns had higher preference rating towards CM compared to breast milk. According to a study on Australian consumers, people preferred the sensory quality and convenience of CM in comparison to soy-milk (Bus & Worsley, 2003). Similar results were also observed by Palacios et al. (2009) where they evaluated the taste and sensory attributes of CM and soy beverages. The study found that with respect to these measures, CM was more preferred than soy beverages. Other variables such as, gender, age, ethnicity, and lactose intolerance do not affect

these results. In the most recent study by Idamokoro, Gunya, and Aliber (2019), people preferred the taste of CM more than GM as GM has a strong smell and distinct taste possibly due to its fatty acid properties.

Although CM has many health benefits and positive effects, it can have negative effects such as cow's milk protein allergy (CMPA), lactose intolerance, and various cancers (Agostoni & Turck, 2011). These are discussed further in the next paragraphs. In general, CM has around 30 – 35 g of proteins per litre. However, some of these proteins can trigger allergic reaction. CMPA is one of the most common causes of food allergy in infants, affecting around one in 50 children (Lifschitz & Szajewska, 2015). The signs and symptoms of CMPA range between mild to severe and can includes the swelling of face, eyes, or lips, hives, vomiting, and gastrointestinal problems (Lam et al., 2008). CMPA is typically outgrown during childhood and adolescence, but some might have it until adulthood or even for the rest of their lives (Hochwallner, Schulmeister, Swoboda, Spitzauer, & Valenta, 2014).

Another negative effect of CM consumption is lactose intolerance. Lactose intolerance occurs when the body is unable to digest lactose, owing to the total or partial absence of an enzyme specialised in breaking down the lactose in milk called lactase (Rangel et al., 2016). Although all milk contains lactose, CM has comparatively higher quantities of it. The signs and symptoms for lactose intolerance include bloating, abdominal pain, diarrhoea, and sometimes vomiting (Mattar, de Campos Mazo, & Carrilho, 2012).

Other than CMPA and lactose intolerance, CM is also linked to some types of cancers. The excessive milk consumption in women is linked with the increased consumption of saturated fatty acid, which is in turn linked to the risk of ovarian cancer (Davoodi, Esmaeili, & Mortazavian, 2013). Women who consume ≥4 servings of total dairy products per day had doubled their risk of having ovarian cancer compared with those who consumed ≤2 servings of dairy products per day (Xu et al., 2007). At the same time, an excessive calcium consumption and the increased milk protein consumption in men may increase the risk of prostate cancer (Preble et al., 2019; Vasconcelos, Santos, Ravasco, & Neves, 2019). Thus, because of these apparent negative health effects of CM, there is a growing demand for alternatives to CM. Consequently, many researches has been performed to find alternatives by focusing on both plant-based and animal-based milk (from different species).

For plant-based milk, there are almond milk, soy milk, oat milk, hemp milk, coconut milk and rice milk. Compared with other plant-based milks, almond milk has low calories. Yet, almond

milk contains good levels of antioxidant and vitamin (Vanga & Raghavan, 2018). In terms of protein, soy milk is the only plant-based milk with protein content (2.92 g/100 g) similar to CM (3.33 g /100g) (McClements, Newman, & McClements, 2019). Oat milk predominantly contains high amount of dietary fibre (2.3%–8.5%). However, unlike CM, oat milk does not contain calcium which is essential for human growth and development (Tallapragada & Rayavarapu, 2019). According to House, Neufeld, and Leson (2010), hemp milk provides protein that is easily digestible. It is one of the few plant-based products that contains the essential amino acids needed for human body. The downside is that hemp milk might not be preferred because of its chalky texture and earthy aftertaste (Callaway, 2004). On the other hand, coconut milk tastes great and it does not contain cholesterol. Nevertheless, coconut milk contains very low amount of protein (0.67 g/100 g) and it is not ideal as an alternative for CM (McClements et al., 2019). Another alternative that may be considered is rice milk. Rice milk is lactose free, making it perfect for someone with lactose intolerance (Lomer, Parkes, & Sanderson, 2008). In addition, rice milk can also act as an alternative for those who are allergic to almonds and soy products.

Based on the nutritional content of plant-based milk stated above, the nutrients content of plant-based milk is still inferior compared with CM. Thus, it might be worth considering milks coming from other animals as alternatives to CM, especially for those with CMPA. Examples for CM's alternatives including milk coming from yak, buffalo, horse, donkey, sheep, goat, and camel. However, considering that some of these animals are not present in NZ, only milk coming from sheep and goat are considered in this literature review.

2.2.2 Goat Milk (GM)

The history of goat milk (GM) had begun 12,000 years ago, where goat was the first animal to be domesticated. For thousands of years, goat has successfully adapted to desert, tropical, and mountainous areas, where other livestock would not be able to thrive (Amills, Capote, & Tosser-Klopp, 2017). Goats are physiologically and morphologically different from sheep and cattle, and they are known to be browsers while sheep and cows are grazers (Sanon, Kaboré-Zoungrana, & Ledin, 2007). As goats are browser, they are good for weed control and prevention of wild-fires (Lovreglio & Ouahiba, 2014). Goats may walk up to 10 km per day foraging, and sheep and cows may walk up to 6.5 km and 5.6 km respectively. In drier areas, goats can survive with much less water intake compared with cows and sheep at 3 days (Haenlein, 2001). Additionally, goats have smaller greenhouse gas (GHG) emission levels compared to, dairy sheep and cattle. According to Opio et al. (2013), the global GHG of dairy

goat and sheep are 20% to 25% those of beef cattle and dairy cattle. For this reason, goats are seen as sustainable livestock, supporting the livelihoods of millions of people and contribute to healthy diets and resilience (FAO, 2020b).

Dairy goats are also known as 'the poor man's cow'. They are strongly associated with low income and food deficit countries where 61% of the world's dairy goat population are found (Pulina et al., 2018). Nevertheless, dairy goat is of significant importance to high-income and developing countries too. For instance, people from South Asia and Sub-Saharan Africa might live on yoghurt and cheeses from SM and GM because they cannot afford CM's product, while people from Mediterranean countries such as Italy, Greece, France, and Spain live on goats and sheep products because it is their tradition and preferences. This proves that consumption of dairy goat products is not necessarily related to underdevelopment or even poverty (Haenlein, 2001).

In 2019, there were approximately 1.05 billion goats in the world (1,045,915,764 head) and 22% of these (217 million) were dairy goats (Miller & Lu, 2019). Despite the large number of the goat population, GM only represented 1.9% of the world's milk production (18.7 Mt) in 2018 (FAOSTAT, 2018b). Even so, there has been a rapid growth in the global dairy goat industry because of the health and nutritive values associated with GM and its products. GM production is expected to increase by 9.7 Mt (+53%) from 18.3 Mt by 2030 (Pulina et al., 2018). According to FAOSTAT (2018a), 56.8% of world's goat are found Asia, predominantly in China, India and Bangladesh; followed by Africa (24.5%), Europe (14.5%), America (4%), and less than 1% in Oceania.

Although the dairy goats industry in Oceania only accounts for $\leq 0.1\%$ of the world's goat population, Oceania produces 9% of the world's goat milk (Miller & Lu, 2019). Despite their smaller body size, goats produce (1.47 times) more milk in comparison with sheep. In 2017, there were an approximately 92 farms with dairy goats in NZ. According to Lopez-Lozano et al. (2017), 72% of dairy goat populations within the country are concentrated mainly in Waikato area, while the remaining 28% are spread throughout the country. In terms of the breed of the dairy goats, the Saanen made up 85% of the total with the rest, 15%, split between; Toggenburg, British Alpine, and Nubian (Solis-ramirez J, 2011).

In NZ, GM products are not as popular as CM products. In fact, GM products are not widely available for public consumption. GM is more likely to be consumed by those who owns dairy goats farm. One of the many reasons for this is because of GM has a distinctive taste. GM

reportedly contains the fatty acid caproic acid, which gives the milk a characteristic "goaty" taste, which grows stronger as the time passes (Kompan & Komprej, 2012). Other than that, the distinctive taste of GM is also felt when the milk is warm (Whetstine & Drake, 2006). Another reason why GM is not as popular as CM could possibly be that GM is expensive. It costs around $17 for a kilogram (kg) of GM solids while it costs $6.75 and $16 per kg of CM and SM solids respectively (Hall, 2019). Additionally, the demand for NZ GM products is higher overseas compared with the demand for GM within the country, especially GM-based infant formula. The main producer of GM products in NZ is Dairy Goat Cooperative and it mainly exports GM powders and infant formulas to 20 different countries around the world (Scholtens, Lopez-Lozano, & Smith, 2017).

Overall, GM's properties were found to be superior to CM. GM has better digestibility and lower allergenic properties compared with CM, making it a suitable alternative, especially for those with CMPA. It has smaller milk fat globules with high level of fatty acids useful for the prevention of fatty liver syndrome making it easier to digest compared with CM (Clark & Mora García, 2017; Silanikove, Leitner, & Merin, 2016). On top of that, short and medium chain fatty acid that are present in GM have been confirmed to be beneficial in several medical treatments. It can control sleeping disorders and digestive problems in children, and reduce cholesterol levels in adults (Haenlein, 2001, 2004).

Numerous studies in the past have attempted to compare the health effects of GM as a substitute for CM. The study from Mack (1952) showed that children consuming GM had increased weight gain, height, and skeletal mineralization compared with children consuming CM. Razafindrakoto et al. (1994) also found that the consumption of GM supported weight gain and fat absorption in undernourished children. Additionally, the nutritional content of GM was also found to be higher than CM. GM has higher oligosaccharides, retinol content, and free amino acids (FAA) compared with CM (Muehlhoff, Bennett, & McMahon, 2013). More information regarding the nutritional content is further discussed in **Section 2.3.**

2.2.3 Sheep Milk (SM)

Sheep were one of the first animals to be domesticated after goats due to their behaviour, manageable size, social nature, and lack of aggression (C.F. Balthazar et al., 2017; Lankin, 1997). The history of sheep domestication goes back to 10,000 years ago. For a long time, sheep were used as a valuable source of nutrients (meat and milk) and clothing material (wool). As a result, different dairy breeds were selected and sheep herds were formed (Barłowska,

Szwajkowska, Litwińczuk, & Król, 2011). In 2018, the worldwide production of SM had reached 10.6 million tonnes, and the top-five producers of SM were China, Turkey, Greece, Syria, and Romania (FAOSTAT, 2018c).

In NZ, there has been an interest in building alternative dairy production systems involving sheep. Sheep dairying is allegedly one of the fastest growing livestock farming opportunity in the country (Griffith, 2015). In NZ, there are approximately 5.6 sheep for every person, making NZ home to 26.7 million of sheep (Statista, 2019). Despite this, not all sheep are suitable for the dairy industry, most are used mainly for the production of meat and wool (Scholtens et al., 2017).

As of 2019, there were 16 dairy sheep companies operating in NZ (Hall, 2019). This is a major increase as there were only five sheep dairy operations back in 2014 (Peterson & Prichard, 2015). The major sheep breeds that are present in NZ are East Friesian (most common in NZ), Lacaune (most common in France), Awassi (most common in Middle East), and Assaf (the crossbred of Awassi and Friesian). In 2019 however, Maui Milk had developed the world-first composite breed called Southern Cross. Southern Cross is the result of the crossbreed between East Friesian, Awassi, Lacaune, and Coopworth (Hall, 2019).

Around the world, SM is considered as a delicacy and its consumption in liquid form is rare. For this reason, countries such as Australia and NZ focus on processing SM into high quality dairy products (Bencini & Pulina, 1997). The exceptionally high levels of fats, proteins, calcium, and conjugated linoleic acid (CLA) in SM makes it an excellent ingredient for cheese productions (Milani & Wendorff, 2011). Several varieties of the most widely consumed cheeses in the world, such as *feta* (Greece), *Pecorino Romano* (Italy), *Ricotta* (Italy), *Roquefort* (France) and *Manchego* (Spain) can be made from SM. Additionally, the same amount of SM could produce more cheeses when compared with CM (Sinanoglou et al., 2015). Besides cheese, SM is also used to produce yoghurt in Greece, as well as gin, vodka, and liqueurs in NZ.

In terms of nutrients value, SM has higher levels of proteins, lipids, minerals, and other essential vitamins in comparison with CM and GM (Park, Juárez, Ramos, & Haenlein, 2007). SM contains twice as many proteins as GM and CM and it is rich in minerals such as calcium, phosphorus, manganese, and magnesium. It also has both saturated fatty acid, and mono- and polyunsaturated fatty acids. According to Rasmussen, Andersen, Jespersen, Mouritsen, and Ditzel (2010), saturated fatty acids such as caproic (C6:0), caprylic (C8:0), and capric (C:10)

acids found in SM can reduce body fat and body weight, while butyric acids (C4:0) commonly found in CM are good to inhibit the growth of human cancer cells. Additionally, the mono- and polyunsaturated fatty acids in SMS are useful in the prevention of cardiovascular diseases and reduce the risk of other non-transmissible chronic diseases (C. F. Balthazar et al., 2016).

SM has smaller fat globules compared with CM and GM, making it more homogenous (Rako, Kalit, & Kalit, 2019). The size and dispersion of fat globules in milk influences its digestibility; the smaller the fat globules, the easier it is to digest and the less likely it is to cause increase cholesterol (Skeaff, Williscroft, Mann, & Chisholm, 2004). According to the study by Masoodi and Shafi (2010) on the alpha casein s1 and s2 proteins in CM, GM, and SM, the protein sequence of SM is similar to that of GM but different from CM. They found that SM has lower allergic sensitization compared with GM and CM. Owing to its high nutrients, SM is also considered to be an alternative for CM, especially for those with CMPA (S. D. Kalyankar, Sarode, Khedkar, Deosarkar, & Rd, 2016). Even so, since the level of lactose in SM is about the same as CM, SM might not be ideal for people with lactose intolerance (Lordan, Tsoupras, Mitra, & Zabetakis, 2018). More information about the composition of SM is further discussed in the next section.

2.3 Nutritional Composition of Cow, Goat, and Sheep Milk

Milk is a complex oil-in-water emulsion containing constituent such as; water, proteins, fat, lactose, and various minerals (Skibiel, Downing, Orr, & Hood, 2013). It also contain a wide variety of bioactive compounds such as peptides, metabolites, nucleotides, immunoglobulins, and other immune proteins (Park, 2009). Milk composition are species specific; this means the milk of mammalian species are designed to supply its offspring with nutrients needed for its growth and survival (Barłowska et al., 2011). The composition of milk varies from species to species and is mainly influenced by the stage of lactation, breed, age, quality of feed, genetics, the length of gestation and dry period, environment, body weight, season, and disease (e.g., mastitis) (Claeys et al., 2014; Jenkins & McGuire, 2006).

Milk composition determines the nutritional value of the mammal's milk. However, it is also the deciding factor on whether or not the milk is suitable as a raw ingredients for manufacturing dairy and other food products, as well as determining the physicochemical and organoleptic properties of these products (Alichanidis, Moatsou, & Polychroniadou, 2016).

Whilst the majority of the world's milk products originate from cows, as discussed in the previous sections, there has been a considerable attention to GM, SM, and other milk types in

the recent years (Pulina et al., 2018). Given the objective of the present work, more focus will be placed on milk from cows, goats, and sheep.

CM, GM, and SM each have different levels of milk constituents, while depending on the parameters, one milk type has been found to be better than another (Clark & Mora García, 2017). As mentioned in the previous sections, the nutritional quality of GM and SM are considered superior to CM. For this reason, GM and SM are sometimes a target of adulteration. Therefore, there is a need to evaluate the nutritional composition of CM, GM, and SM separately and identify their difference in literature. In this section, the gross composition of cow, goat and sheep milk can be found in **Table 2.1**. The values of the nutritional composition given in this section are presented in a range, rather than the absolute value.

Table 2.1. Gross composition of cow, goat and sheep milk[a]

Parameter	**Cow**	**Goat**	**Sheep**
Water (%)	87.2 - 87.8	87	80.7 - 81.6
Protein (%)	3.2 - 4.0	2.9 - 3.6	5.2 - 6.6
Fat (%)	3.9 - 5.4	3.5 - 4.5	6 - 9.3
Lactose (%)	4.6 - 4.9	3.2 - 4.4	4.2 - 5.7
Ash (%)	0.7 - 0.8	0.9	0.87 - 0.97
Dry matter (g / kg)	105 - 137	119 - 163	152 - 193
Fat globule diameter (µm)	0.92 - 15.75	0.73 - 8.58	0.4 - 6.68
Micelle diameter (nm)	180	260	193
Energy (kcal/kg)	590 - 701	580 - 740	930 - 1080

[a]As adapted from: (Jenness, 1980),(Barłowska et al., 2011),(Merlin Junior et al., 2015),(Getaneh, Mebrat, Wubie, & Kendie, 2016), (Alichanidis et al., 2016),(Osthoff, 2016),(Wendorff & Haenlein, 2017),and (Burrow, Young, McConnell, Carne, & Bekhit, 2018).

Further detailed information on individual components of milk such as proteins, carbohydrates, fats, minerals, and vitamins can be found in the next sections.

2.3.1 Milk Protein

As one of the most studied food proteins, milk proteins are complex and highly versatile. These properties have made it susceptible to many processing conditions in the food and dairy industries (Jenkins & McGuire, 2006). Depending on the type of dairy products, milk proteins have different roles (Andiç & Boran, 2015). Their key functions include gelation, emulsification, and foaming. Additionally, milk proteins are also responsible for the flavour of the fluid milk (Schiano, Harwood, & Drake, 2017). Milk proteins are made of heterogeneous

groups that occur at different amounts and these can be divided into two categories: casein and whey. Other than that, milk also has NPN (Non-Protein Nitrogen) constituents, which are nitrogenous compound that can be converted into by microbes present in the ruminant stomach (Amha, 2015). More information regarding casein, whey, and NPN can be found in the following sections.

2.3.1.1 Casein

Casein is the main protein fractions in ruminant milk. Casein is highly stable, non-toxic, and relatively inexpensive (O'Kennedy, 2011). Casein will precipitate in milk when the pH is near 4.6 at 20°C; while the rest of the fraction, whey protein or serum, is soluble under similar conditions (Dalgleish, 1982). Casein fractions are not homogeneous, they can be distinguished into four fractions: αS1-casein, αS2-casein, ß-casein, and κ–casein (Farrell et al., 2004). Each of the variant's primary structure is determined genetically, and together they differ from each other only by amino acid residues. The proportion of casein protein fractions are not equal in different milk, and in some cases, one or more caseins maybe absent depending on the ruminant species. The alpha (αS1 and αS2) caseins and beta caseins are known to be calcium sensitive. When they bind into calcium, caseins will precipitate at high concentration. On the other hand, kappa caseins are known as the calcium-insensitive casein used to stabilize micelles (Kawasaki, Lafont, & Sire, 2011). Micelles characteristics are different regarding the size, mineralization, and hydration. The casein micelles in CM are more hydrated compared with GM and SM that have higher mineralization levels (C.F. Balthazar et al., 2017).

Table 2.2. Casein fractions of cow, goat, and sheep milk[a,b]

Casein Fractions	Cow	Goat	Sheep
αS1-casein (%)	40	8	6.7
αS2-casein (%)	10	18	22.8
ß-casein (%)	45	55	61.6
κ–casein (%)	5	19	8.9

[a]Sourced from (J. R. Brown, Law, & Knight, 1995),(Farrell et al., 2004),and(Selvaggi, Laudadio, Dario, & Tufarelli, 2014)

[b]Values displayed in the table should not be seen as an absolute value, rather viewed as an approximation

2.3.1.2 Whey Protein

Whey protein in milk consists of non-casein proteins that remain soluble when casein precipitates at pH 4.6 at 20°C (Farrell et al., 2004). Whey proteins are generally globular proteins with many secondary and tertiary structures (Alichanidis et al., 2016). In milk, it presents as very small aggregates. When denatured by heat, whey protein could triggers hydrophobic interaction with other proteins and form a protein gel (Foegeding, Davis, Doucet,

& McGuffey, 2002). The major constituents of whey proteins are α-Lactalbumin, ß-Lactoglobulin, immunoglobulins, serum albumin, and other minor proteins.

α-Lactalbumin is present in the milk of all mammals and it has a specific function in breaking down the lactose in milk (Park et al., 2007). ß-Lactoglobulin is the major whey protein present in mammal's milk (excluding humans and camels) that possess both anti-carcinogenic and antiviral activities (Davoodi et al., 2016). Immunoglobulins are normally found in colostrum, and function to protect neonates against specific pathogens (Ulfman, Leusen, Savelkoul, Warner, & van Neerven, 2018). Serum albumin in whey is a major protein that is found in blood and occurs in all part of the body. There is no known functional role of this protein in milk.

Different species have different percentage of whey proteins. According to Potočnik, Gantner, Kuterovac, and Angela (2011), SM had the highest whey protein fractions followed by goat and cow. Rafiq et al. (2016) discovered similar findings and reported that whey proteins in SM are more prone to heating compared with CM. The whey protein in SM is said to have better foam stability, gel strength, and foam overrun when compared with GM and CM.

Table 2.3. Whey protein constituents in cow, goat, and sheep milk[a-b]

Whey Fractions	Cow	Goat	Sheep
α-Lactalbumin (%)	16.2	21.4	10.8
ß-Lactoglobulin (%)	59.3	54.2	61.1
Immunoglobulin (%)	15	11.5	20
Serum Albumin (%)	9.5	12.8	8.1

[a]Sourced from (Borková & Snášelová, 2005), (Potočnik et al., 2011),(Rafiq et al., 2016), and (Taj Khan et al., 2019)

[b]Values displayed in the table should not be seen as an absolute value, rather viewed as an approximation

2.3.1.3 Non-Protein Nitrogen (NPN)

Non-protein nitrogen (NPN) is defined as a small concentration of nitrogenous substances that are present in milk serum. These substances include urea, free amino acids, small peptides, biuret, ammonia, uric acid, and others (Floris, Lambers, Alting, & Kiers, 2010). The amount of NPN available is different across species. The NPN fractions in sheep are around 5-6.8% of the total nitrogen present in milk, while the NPN fractions in cows and goats are around 5% and 5.8% respectively (DePeters & Ferguson, 1992; Ramos & Juarez, 2011). From this, around 20-75% of the NPN fractions are mainly urea, which are related to the protein and energy supply of the milk (Floris et al., 2010). In general, the NPN fractions in milk could be determined with the Kjehdahl method (Ruska & Jonkus, 2014).

Other than urea, the rest of the NPN fractions are made of free amino acids (FAA) and protein-bound amino acids. FAA is more easily absorbed, and they represent around 10-20% of NPN in CM, 9-10.5% in GM, and 16% in SM (Park et al., 2007). The FAA in milk represents both essential and non-essential amino acids occurring at different concentrations (**Table 2.4**). In general, the concentration of FAA in all kinds of milk indicates the quality of the milk, and this in turn is affected by the lactation stage. High levels of FAA indicate poor quality of milk and it usually occurs during the early or late lactation stage of the mammals (McDermott et al., 2016). Aside from quality, several amino acids are responsible for digestibility, cheese-making properties, and the flavour of the milk (Haenlein, 2004). Additionally, the characterization of amino acids is found to be useful as a detection method of milk adulteration (Tripaldi, Martillotti, & Terramoccia, 1998).

Table 2.4. Free amino acids composition (mg/kg) in the proteins of cow, goat, and sheep milk[a,b]

FAA content (mg/kg)	**Cow**	**Goat**	**Sheep**
Essential amino acids			
Isoleucine	0.3 - 1.4	0.3 - 2.2	0.3
Leucine	0.4 - 2.9	0.3 - 2.7	0.5
Lysine	2.2 - 2.8	3.2 - 5.5	2.6
Methionine	0 - 0.6	1.4	0.3
Phenylalanine	0.5 - 1.6	2	0.3
Threonine	0 - 1.5	0 - 3.3	5
Tyrosine	0.05 - 1.5	1.5 - 4.5	1.5
Valine	0.6 - 6.7	5.8 - 6.3	1.3
Non-essential amino acids			
Alanine	1.0 - 3.4	1.6 - 7.7	5.3
Arginine	0.9 - 1.7	1.9 - 13.5	3.7
Aspartic acid	1.2 -2.6	1.3 - 5.2	1.9
Glutamic acid	7.7 - 17.2	25 - 43.5	28.4
Glutamine	1.6 - 1.8	10.1 - 27.4	10.7
Glycine	0.6 - 6.6	21.8 - 32.4	11.6
Proline	0.5 - 3.2	1.6 - 1.8	0
Serine	0 - 2.4	0 - 9.5	0.3

[a]Sourced from (Rassin, Sturman, & Guall, 1978),(Fenyvessy, Sirokman, & Varro, 1991), (Jandal, 1996), and (Tripaldi et al., 1998)

[b]Values displayed on the table should not be seen as an absolute value, rather viewed as an approximation

2.3.2 Milk Sugars

Milk harbours many bioactive components including naturally occurring sugars. These natural sugars are the main carbohydrate in milk that is responsible for its light and sweet taste (Gambelli, 2017). Milk carbohydrates are essential for a healthy diet. In the human body, they break down into simple sugars, which are later converted into energy used to support physical activity and bodily function (Mozaffarian, Hao, Rimm, Willett, & Hu, 2011). There are

different kinds of carbohydrate present in milk including, lactose, galactose, glucose, and other oligosaccharides.

2.3.2.1 Lactose

Lactose is primarily found in dairy products as the naturally occurring sugars in milk. As a disaccharide made of glucose and galactose bonded together, lactose acts as the source of energy in milk (Johnson & Conforti, 2003). By weight, lactose makes around 2-8% of the total weight of milk and accounts for around 40% of the milk's caloric value (Thorning et al., 2016). Lactose has lower solubility compared with other disaccharides. For this reason, lactose is less sweet compared with sucrose, fructose, and glucose (McCain, Kaliappan, & Drake, 2018).

The amount of lactose present in each dairy product varies greatly depending on the lactose hydrolysis and the mammalian species. For example, the amount of lactose present in the spray dried milk is generally higher than the fluid milk owing to the water removal (Nijdam & Langrish, 2006). For mammalian species, the lactose present in CM, GM, and SM are different based on species type and the lactation stage. As the lactation stage on mammal progresses, the lactose content decreases, while the mineral in milk increases (Dominguez-Salas, Galiè, Omore, Omosa, & Ouma, 2019). The details on the lactose content in CM, GM, and SM can be seen in **Table 2.1**.

2.3.2.2 Oligosaccharides

Unlike lactose that is only present in mammalian dairy products, oligosaccharides are present in all kinds of milk including plant-based milk. These compounds are believed to have wide range of nutritional benefits, especially in the gut health of neonates (Urashima & Taufik, 2010). Different milk oligosaccharides might have a complicated structure even though most of them share the same lactose core consisting galactose and glucose linked with β 1,4-linkage (Lange et al., 2014). These oligosaccharides are produced in mammary glands, and typically contain 3 to 10 branched or linear monosaccharides, with lactose, N-acetylneuraminic acid, and galactosamine on their reducing end (Oliveira, Wilbey, Grandison, & Roseiro, 2015).

Even though oligosaccharides present in milk from all kinds of species, human milk has higher oligosaccharides (5-16 g/L) compared to animal milk (Kunz, Rudloff, Baier, Klein, & Strobel, 2000). Amongst all animal milks, GM has a quite high milk oligosaccharides with 0.23-0.3 g/L, followed by cow with 0.03-0.09 g/L, and sheep with 0.02-0.04 g/L (Abd El-Salam & El-Shibiny, 2011).

2.3.3 Milk Fats

Milk fat is one of the most important components of the nutritional and quality of the milk products, where it affects the texture, mouthfeel, nutritional value, and mechanical properties of dairy products (Ramel & Marangoni, 2019). They are considered as a good source of energy and a reliable carrier of fat-soluble vitamins present in milk (vitamin A, D, E , and K) as well as β-carotene (Gómez-Cortés, Juárez, & de la Fuente, 2018). Additionally, milk fat is the main ingredients in many food products such as: butter, cheese, cream, and ice-cream.

Milk fat is the most variable constituent in milk, and their level varies between and within milching species. The continuous milking process affects the level of fat present in milk, where the milk with lowest fat concentration is drawn in the beginning and milk with the highest fat being drawn in the end (Bernard & Tao, 2019).

Lipids in milk are mainly present as globules in oil-in-water emulsion. Milk fat globules are formed by the endoplasmic reticulum located in the epithelial cells and coated by milk-fat globule membrane (MFGM) that are rich in proteins, cholesterol, glycoproteins, phospholipids, and other polar lipids (Kompan & Komprej, 2012; Månsson, 2008). The presence of MFGM is important in milk, as they are needed to prevent lipid degradation and stabilize the enclosed fat against fusion and coalescence. The detailed composition and structure of MFGM varies widely between species and it has a significant effect in the physical stability of the milk and digestibility (Gantner, Mijic, Baban, Škrtić, & Turalija, 2015).

As mentioned previously, the sizes of milk fat globules were found to have negative correlation with digestibility (Lopez, Cauty, & Guyomarc'h, 2019). The larger the size of fat globules, the harder it is for the milk to be digested. Additionally, MFGM are less stable and they have decreased resistance to coalescence and deformation under chemical pressure (Claeys et al., 2014). Thus, larger fat globules are more prone to disruption during processing. Even so, milk containing larger fat globules is linked alongside milk with high-fat content instead of milk with a lower fat content (Ménard et al., 2010). Amongst CM, GM and SM; CM has the largest fat globules followed by GM and SM (Barłowska et al., 2011).

2.3.3.1 Milk Fatty Acids Profile

Fats are made from individual molecules of fatty acids (FAs) attached to a 3-carbon backbone called glycerol. Triacylglycerol (triglyceride) is the most common type of fat derived from plant and animal sources (Lichtenstein, 2013). The milk fat triacylglycerols can be synthesised

from more than 400 different fatty acids, making it the most complex fat out of all-natural fats (Alichanidis et al., 2016).

Generally, different lengths of FAs have different esterification positions. The short chain FAs such as butyric (C4:0) and caproic acid (C6:0) are typically almost fully esterified at *sn*-3, while the medium FAs such as caprylic (C8:0), capric (C10:0), lauric (C12:0), myristic (C14:0) as well as the long chain palmitic acid (C16:0) are esterified at positions *sn*-1 and *sn*-2. The other long FAs such as stearic acid (C18:0) are selectively esterified at *sn*-1, though oleic acid (C18:1) showed preferences for positions such as *sn*-1 and *sn*-3 (Månsson, 2008; Parodi, 2004).

Regardless of species, milk fat globules are mainly composed of saturated fatty acids (SFA), unsaturated fatty acids including, monounsaturated (MUFA) and polyunsaturated fatty acid (PUFA), *cis, trans,* and conjugated fatty acids. Depending on the type and amount of fatty acid in each milk, they each have either positive or negative effects on human health.

Saturated fatty acids (SFA) are the primary fat component of the human diet, commonly found in animal products. SFA accounts for around 60-70% of fatty acids present in ruminant milk and they have various chain lengths (Markiewicz-Keszycka, Czyżak-Runowska, Lipińska, & Wójtowski, 2013). According to Månsson (2008), the concentration of SFA in milk is reported to be lowest in the summer and highest in winter, while the concentration of unsaturated FA shows the opposite pattern, being highest in the summer and lowest in winter. From all milks, palmitic acid (C16:0) is said to be the highest from quantity viewpoint and accounts for around 30% of the total fatty acids, followed by stearic acid (C18:0), and myristic acids (C14:0) which can made up around 11 and 12 % of the milk weight from ruminant species (Loften et al., 2014; Piantoni, Lock, & Allen, 2013). GM and SM are rich in medium FA compared with CM. Excessive amount of SFA in the human diet could lead to obesity, atherosclerosis, and other chronic diseases.

Polyunsaturated fatty acids (PUFA) are among the healthy fats, along with monounsaturated fatty acids (MUFA) (Ander, Dupasquier, Prociuk, & Pierce, 2003). In ruminant milks, PUFA accounts for about 5-9% of the total FAs present. PUFAs are primarily consisted of omega-3 and omega-6 fatty acids. In cow, goat, and sheep milk, omega-3 fatty acids including eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) are found in trace amounts. The other omega-3 fatty acids, alpha-linoleic acid (ALA) is found in higher concentration in GM and SM compared to CM (Markiewicz-Keszycka et al., 2013). On the other hand, a type of omega-6 fatty acid, conjugated linoleic acid (CLA) accounts for around 15-20% of PUFAs

present in ruminant milk. SM is reported to have the highest CLA content followed by CM, and GM (Abd El-Salam & El-Shibiny, 2011). Overall, all types of PUFA in human diets are perceived as a beneficial dietary intervention for the treatment and prevention of high blood pressure, atherosclerosis and other CVD (Ander et al., 2003).

Monounsaturated fatty acids (MUFA) are classified as FAs containing single double bond with *cis* and *trans* configurations. In cis-configuration, the hydrogen atoms are positioned on the same side as the double bond; while on the trans-configuration, the hydrogen atoms and the double bonds are positioned on the opposite sides (Schwingshackl & Hoffmann, 2012). Unlike SFA, MUFA does not cause accumulation of cholesterol. Different from PUFAs, MUFAs are relatively stable and do not readily become rancid (DiNicolantonio & O'Keefe, 2017). The level of MUFAs present in cow, sheep and goat milk fat are similar and they range between 20 and 35% of the total FAs. In the human diet, MUFAs are said to be able to reduce inflammation and improve insulin sensitivity (Cruz-Teno et al., 2012; Paniagua et al., 2007). The share of fatty acids profile in the three milk types can be seen on **Table 2.5**.

Table 2.5. Fatty Acids Profile (g/ 100g) in cow, goat, and sheep milk[a,b]

Fatty Acids Constituent (g/100g)	**Cow**	**Goat**	**Sheep**
C4:0; butyric	2.9	2.0	2.6
C6:0; caproic	2.0	2.8	1.9
C8:0; caprylic	1.4	2.9	1.9
C10:0; capric	3.0	9.6	6.6
C12:0; lauric	3.6	4.5	4
C14:0; myristic	11	9.8	10.2
C16:0; palmitic	29	25	25
C18:0; stearic	11	8.9	8.9
C18:1 *cis*-9; oleic	22.4	18.7	20.2
C18:2 *cis*-9, *cis*-12; linoleic	2.6	2.3	2.3
C18:2 *cis*-9, trans-11; CLA	0.57	0.45	0.76
ω-3	0.53	0.44	1.31
ω-6	0.5	1.72	2.97
SFA	72	72	69
MUFA	25	30	27
PUFA	4	4	5

[a]Sourced from (Månsson, 2008),(Butler, Stergiadis, Seal, Eyre, & Leifert, 2011),(Barłowska et al., 2011) , (Markiewicz-Keszycka et al., 2013), (Claeys et al., 2014), and (Alichanidis et al., 2016).

[b]Values displayed on the table are the average value calculated form different data set. It should not be seen as an absolute value, rather viewed as an approximation

2.3.4 Minerals

Milk is an important source for mineral components that, are essential for humans to function. Depending on their concentration, these minerals are classified as macro-elements (calcium,

phosphate, magnesium, sodium, and potassium) and also in micro- or trace elements (zinc, cooper, iron, manganese, selenium, etc.) (Bilandžić et al., 2014). From all the mineral compounds present in milk, calcium and phosphorous are the two main components, known to be crucial for strong healthy bones and energy production in the body (Pietrzak-Fiećko & Kamelska-Sadowska, 2020). The presence of all other salts are important for the technological properties in milk and its nutritional viewpoint, where they contribute to maintaining the milk buffering capacity, pH, ionic strength and milk osmotic pressure (Lucey & Horne, 2009).

Amongst the milk of the common ruminants, SM has the highest concentration of calcium, phosphorous, and magnesium. CM has the highest share of sodium, copper, and iron, while GM has the highest share of potassium and lowest concentration of iron and copper (**Table 2.6**). Despite having a low concentration of iron, the iron contents in GM are more bioavailable than CM due to the presence of nucleotides that, increase the intestinal absorption of iron (Pietrzak-Fiećko & Kamelska-Sadowska, 2020; Raynal-Ljutovac, Lagriffoul, Paccard, Guillet, & Chilliard, 2008)

Table 2.6. Minerals concentration in cow, goat, and sheep milk[a,b]

Minerals	**Cow**	**Goat**	**Sheep**
mg / 100g			
Calcium	122	134	195
Phosphorus	119	121	141
Potassium	152	185	138
Magnesium	12	16	19
Sodium	49	41	39
µg / 100g			
Zinc	6.2	6.9	5.8
Iron	8	6	7
Copper	6	4	1
Manganese	6	8	7
Selenium	1.8	1.6	1.7

[a]Sourced from (Y. W. Park et al., 2007),(Raynal-Ljutovac et al., 2008),(Barłowska et al., 2011),(Claeys et al., 2014) , and (Pietrzak-Fiećko & Kamelska-Sadowska, 2020).

[b]Values displayed on the table are the average value calculated form different data set. It should not be seen as an absolute value, rather viewed as an approximation

2.3.5 Vitamins

Vitamins are a group of chemical compounds essential in a small amount for cell metabolism and functions (Ward, 2014). Since vitamins cannot be synthesized by humans, they need to be present in the diet (Graulet, 2014). Milk is an excellent source of vitamins as it contains all 13 known vitamins, including the water-soluble and the fat-soluble vitamins.

The water soluble vitamins present in milk includes vitamin B complex (thiamine (B1), riboflavin (B2), niacin (B3), pantothenic acid (B5), pyridoxine (B6), biotin (B7), folate (B9), and cyanocobalamin (B12)) and ascorbic acid (Vitamin C); while the fat soluble vitamins includes; vitamins A, D, E , and K (Graulet & Girard, 2017).

Since all vitamins have their own role in the human body system, it is especially important for humans to have an adequate intake of vitamins. Inadequate intake of vitamins induces serious health hazards; for example, the lack of vitamin A could lead to blindness (Meyers, Hellwig, & Otten, 2006). More detailed information about the dietary reference intake for each vitamin, their role and consequences of deficiencies can be found in **Table 2.7**.

Table 2.7. Dietary reference intake for vitamins, its main role, and consequences of deficiencies in human health[a]

Vitamins	**Daily Reference Intake (µg/day)**		**Main role**	**Consequences of deficiencies**
Fat-soluble Vitamins	Children	Adults		
Retinol (A)	300 - 500	600 - 900	Form and maintain immune function, vision, and reproduction	Blindness, dry skin, xerophthalmia
Cholecalciferol (D)	5	5 - 15	Maintain bone health, calcium, and phosphate homeostasis	Multiple sclerosis, cancers, rickets
Tocopherol (E)	4000 - 7000	11000 -15000	Antioxidant, maintain healthy skin	Muscle weakness, vision problems, weak immune system
Phylloquinone (K)	2 - 55	60 - 120	Blood coagulation	Osteoporosis, CVD, bleeding
Water-soluble Vitamins				
Thiamine (B1)	200 - 600	900 - 1200	Process carbohydrate, fats, and protein	Beriberi, short term memory loss, weight loss
Riboflavin (B2)	300 - 600	900 - 1300	Growth and maintain good health	Skin disorders, sore throat, hyperaemia, cheilosis
Niacin (B3)	2000 - 8000	12000- 16000	Lower cholesterol, prevent CVD	Pellagra, dementia, diarrhoea
Pantothenic acid (B5)	1700 - 3000	4000 - 5000	Synthesize and metabolize fats and proteins	Irritability, numbness and tingling of hands and feet,
Pyridoxine (B6)	100 - 600	1000 - 1700	Heme synthesis and amino acid metabolism	Convulsions, nervous system impairment, skin changes
Biotin (B7)	5 - 12	20 - 30	Liver metabolism, strengthen hair and nails	Conjunctivitis, central nervous system disorder, dermatitis, depression
Folate (B9)	65 - 200	300 - 400	convert carbohydrates into glucose (food into energy), produce red blood cells, and DNA synthesis	Mouth sores, megaloblastic anaemia
Cobalamin (B12)	0.4 - 1.2	1.8 - 2.4	produce genetic materials in cells, prevent anaemia	Anaemia, fatigue, damage to nervous system
Ascorbic Acid (C)	15000 - 50000	45000 - 90000	growth, development and repair body tissues, wound healing, protect against CVD	Scurvy, anaemia, fatigue, rashes

[a]Sourced from (WHO, 2004) ,(Ball, 2008), (Graulet, Martin, Agabriel, & Girard, 2013)

Since milk is a crucial source of vitamins, the concentration of vitamins present in milk is highly variable depending on several factors. The factors include, season, breed, maternal diet, geographical location and its vitamin status (Alichanidis et al., 2016; Thompson, Henry, & Kon, 1964). Even so, the water-soluble vitamins are more responsive towards maternal diet than the fat-soluble vitamins. The amount of fat-soluble vitamins in milk depends on the fat content of the product (Górska-Warsewicz et al., 2019). Amongst the milk from the common ruminant, sheep milk is richer in most of the vitamins compared with the milk from cow and goat (**Table 2.8**).

Table 2.8. Vitamin content of cow, goat, and sheep milks[a,b]

Vitamins	**Cow**	**Goat**	**Sheep**
Fat-soluble (µg / 100 g)			
Retinol (A)	39	51	63
Cholecalciferol (D)	0.18	0.07	0.18
Tocopherol (E)	122	58	120
Phylloquinone (K)	trace amount	trace amount	trace amount
Water-soluble (µg / 100 g)			
Thiamine (B1)	45	68	80
Riboflavine (B2)	160	210	376
Niacin (B3)	80	264	420
Pantothenic acid (B5)	350	320	408
Pyridoxine (B6)	47	42	72
Biotin (B7)	2.8	2.4	3.2
Folate (B9)	5.5	1	4.6
Cobalamin (B12)	0.4	0.16	0.7
Vitamin C (mg / 100g)	0.94	1.4	4.16

[a]Sourced from (Y. W. Park et al., 2007), (Raynal-Ljutovac et al., 2008), (Barłowska et al., 2011), and (Alichanidis et al., 2016)

[b]Values displayed on the table are the average value calculated form different data set. It should not be seen as an absolute value, rather viewed as an approximation

2.4 Milk Adulteration

Although milk is considered optimum source for protein, carbohydrate, fat, vitamins, and minerals, milk is easy to be adulterated (Brantsæter, Olafsdottir, Forsum, Olsen, & Thorsdottir, 2012). In fact, milk is one of the most likely food items at risk of adulteration (following olive oil) (Moore et al., 2012). The adulteration of milk will result in the reduction of the milk quality and, depending on the adulterants, it can also cause adverse risks on human health (Poonia et al., 2016). There are many reasons behind milk adulteration. It can be either intentional or accidental.

Intentional adulteration happens when compounds are added into milk intentionally, with knowledge to earn a profit or hide quality defects (Spink, 2014). On the other hand, the accidental adulteration could be attributed to carelessness, ignorance, and the lack of proper facilities and hygiene during food processing (Bansal, Singh, Mangal, Mangal, & Kumar, 2017).

Even though most adulterations are done for economic benefits, their impact could cause serious problems for enterprises, farmers or producers, and consumers (Handford, Campbell, & Elliott, 2016). For instance, when the distributed milk is known to have been adulterated, the enterprise would be impacted with the loss of consumer trust on the product, the cost of product recalls, complaints, and other cost related to the replacement of and compensation for the product (Ayza & Belete, 2015). All these will also result in the lack of product acceptance and a decrease in product demand, affecting farmers and producers (Qian, Guo, Guo, & Wu, 2011). Moreover, depending on the adulterants, such fraudulent activities can cause severe health effects including kidney failure, diarrhoea, abdominal pain, headache, gastroenteritis, paralysis and sometimes leading to death (Handford et al., 2016).

2.4.1 Common Milk Adulterants and Public Health Risks

Since milk adulteration is very common in some developing countries, it is most important for consumers to know the different type of adulterants and their effect on health (Salih & Yang, 2017). Common adulterants added into milk include water, vegetable protein, whey, and milk from another species. When these adulterants are added, consumers usually cannot tell whether the milk has been tainted (Fischer, Schilter, Tritscher, & Stadler, 2015). Even though, adulterants such as water, vegetable protein, and whey, do not pose severe health risk, there are adulterants that are too harmful to be overlooked (Azad & Ahmed, 2016).

Some of the major milk adulterants with adverse health effect include formalin, salicylic acid, anionic detergent, hydrogen peroxide, melamine, urea, caustic soda, and benzoic acid. These compounds are added to milk with the intention to make the milk quality and its nutritional properties seem higher (Nascimento et al., 2017). More detailed information about the common milk adulterants and their public health risks can be seen in **Table 2.9**.

Table 2.9. Common milk adulterants, uses and their public health risks

Adulterants	Type of Adulterants	Functions	Public Health Risks	People Who are at Risk	References
Urea	To increase milk mass and nutritional content	Artificially increase the level of protein in milk and blended with other ingredients to produce synthetic milk	Kidney malfunction, Indigestion, Acid, Ulcers, Damage to intestinal tract	Infant, Young Children, Pregnant Women, Elderly	(Finete, Gouvêa, de Carvalho Marques, & Netto, 2013; Paradkar, Singhal, & Kulkarni, 2000)
Anionic Detergent	To improve the sensory attributes of milk	Enhance the milk's cosmetic nature as the foamy appearance diminishes when it is diluted with water	Hypotension, Gastrointestinal problems, Cancers, Respiratory irritation	Individuals from developing countries with poor food safety standard	(Ahirwar, Harilal, Srihari, & Pandey, 2015; Azad & Ahmed, 2016)
Hydrogen Peroxide	To preserve milk	Reducing the electricity cost by extending the shelf life of milk without refrigeration	Vomiting, Nausea, Lethargy, Gastritis	Individuals from developing countries that is lacking in cooling facilities	(Bansal et al., 2017)
Caustic Soda (Sodium Hydroxide)			Skin irritation, Vomiting, Severe cases could lead to damage of oesophagus	Infant, Young Children	(Aiello, Pizzolongo, Manzo, & Romano, 2019)
Salicylic Acid			Diarrhoea, Gastric irritation, Bleeding, Severe cases could lead to poisoning and even death	Individuals from developing countries that are lacking in cooling facilities	(P. Singh & Gandhi, 2015)
Formalin		Reducing the electricity cost by extending the shelf life of milk for a very long time at room temperature	Abdominal pain, Cancer, Vomiting, Liver and Kidney Damage, Severe cases can lead to blindness		(Mabood et al., 2016; X. Tang et al., 2009)
Benzoic Acid		Reducing cost by increasing the shelf life of milk	Asthma, Urinary tract problems, and behaviour disorder in children	Infants, Children, Individuals who are allergy sufferers	(Barham, Khaskheli, Soomro, & Nizamani, 2014)

Table 2.9. Common milk adulterants, uses and their public health risks (continued)

Adulterants	Type of Adulterants	Functions	Public Health Risks	People Who are at Risk	References
Vegetable oil	To increase milk mass and nutritional content	Milk fat is replaced by vegetable oil for economic gain	Some vegetable oil might contain nut oils, and if the consumer is allergic to nuts, this could trigger a severe allergic reaction including hives, diarrhoea, coughing, and wheezing. In some cases, it could be life threatening for individuals	Individual with nut allergy	(P. Singh & Gandhi, 2015)
Ammonium Sulphate	To increase milk mass and nutritional content	Artificially increase the level of protein in milk	Nausea, Vomiting, Diarrhoea, Skin reactions, Gastrointestinal problems	Infant, Young Children	(Ahirwar et al., 2015)
Melamine			Kidney stone, Bladder cancer, Acute renal failure, Hypertension, In worst cases could led to death	Infant, Young Children	(Chan, Griffiths, & Chan, 2008; Xiu & Klein, 2010)

2.4.2 Milk Adulteration Incidents

As there are many types of adulterants for milk, the history of milk adulteration had gone back 170 years ago (Atkins, 1991). The first known case of milk adulteration was the Swill milk scandal in New York in the 1850s, where it caused the death of 8,000 infants. In 1858, a journalist had discovered that these infants given swill milk coming from cows fed on distilleries waste. These cows were fed on the swill, mash, and run-off from whiskey distilleries. They were kept in horrible conditions to the point where they often stood in their manure, were covered with cold sores and were suffering from a range of diseases. As a result, the cows produced milk that was in an unnatural bluish colour. Hence, several components such as starch, flour, egg, and plaster of Paris were added to the milk to make it appears thick and white. However, this was a health hazard and resulted in the death of 8,000 infants (Wilson, 2008).

Another known milk scandal was the Morinaga milk arsenic poisoning incident in Japan in 1955. Arsenic was inadvertently added to the milk through the industrial grade of monosodium phosphate, which was used as an emulsifier and thickening agent in milk. This incident resulted

in the deaths of over 100 infants and chronic health effects in 13,000 people (Dakeishi, Murata, & Grandjean, 2006).

The most significant milk adulteration incident happened in China, 2008. The infant formulas manufactured by the Chinese company, Sanlu Co., Ltd were found to be adulterated with melamine, resulting, in the deaths of six infants and the illness of 300,000 young children and infants (Xin & Stone, 2008). Melamine is high in nitrogen and is relatively cheap. Because it is high in nitrogen, the addition of melamine in milk artificially increases the apparent protein content as measured by the standard test (Gossner et al., 2009). Nonetheless, there are no approved reasons for melamine to be added to food in the world (Reinberg, 2008). In fact, the addition of melamine into the formula was causing infants to develop kidney stones, which, when left untreated, could cause renal failures and even deaths (Xiu & Klein, 2010). This incident brought the public's attention to milk adulteration. It sparked many controversies including the involvement of Fonterra, NZ's largest dairy company (see next section).

2.4.2.1 New Zealand's Involvement in Chinese Milk Scandal

To date, there has been no milk adulteration case in NZ. However, Fonterra, NZ's largest dairy company was involved in the 2008 Chinese milk scandal (Coonan, 2013). Fonterra owned 43% shares of Sanlu Co., Ltd; the company behind the adulteration of infant formula (Chan et al., 2008). At the time of the incident, Fonterra reportedly knew about the addition of the melamine one month before the case went public and claimed to have pushed for product recall. Though there was an immediate recall during some trade, Sanlu denied the request for official product recall (Fu & Nicoll, 2016). The official total recall of the product only happened after the NZ prime minister alerted the Chinese government.

In the aftermath of the event, Sanlu went bankrupt. A few local government officials were forced to resign, three company executives from Sanlu got life imprisonment, two received 15 years jail term, one got suspended death penalty, and two were executed (Barboza, 2009). Several milk dealers and suppliers were also charged and arrested over selling the melamine. As Fonterra owned a large share of the company, it also suffered a huge financial loses (Scott, Bowden, & Rowarth, 2013).

2.4.3 Risk of Adulteration of High Value New Zealand Dairy Products

The dairy industry in NZ acts as one of the major sectors of the country's economy. Globally, NZ dairy products have a clean green image from the high degree of food safety level in the country (Ballingall & Pambudi, 2017). In NZ, the Food Act 2014 is the primary legislation for

governing food safety, managed by the Ministry of Primary Industries (MPI, 2018). Based on the provisions in the Food Act, milk adulteration is illegal in NZ. In addition to the Food Act 2014, all the milk and dairy products in the nation are strictly monitored through the National Contaminants Control Programme (NCCP) to ensure that there are no contaminants in the products (MPI, 2013). Thus, the likelihood of milk being adulterated in NZ is very low.

Even so, 95% of NZ dairy products are exported to more than 150 countries in the world (OEC, 2017). The milk exported to other countries includes milk and milk-based products from cows, goats, and sheep (Griffith, 2015; Scholtens et al., 2017). As a highly valued product in the public eye, NZ milk is more expensive compared with milk from other countries (Shadbolt & Apparao, 2016). Therefore, to gain more profits the overseas distributor could alter the properties of NZ milk by adding adulterants. As a result, it is important to have a robust analytical method that can effectively detect the presence of adulterants in milk.

2.5 Methods to Detect Milk Adulteration

There are many ways to detect adulterants in milk. Depending on the types of the adulterants, both qualitative and quantitative detection methods can be performed on the milk products.

2.5.1 Qualitative Detection of Milk Adulteration

The qualitative methods also known as the traditional method to detect adulteration are relatively fast and simple to perform (Reddy, Venkatesh, & Reddy, 2017). These are usually colour based chemical reactions, that can be performed in any biosafety level 1 laboratory (BSL-1) subject to availability of reagent (Azad & Ahmed, 2016). Examples of the adulterants that can be detected through rapid qualitative method can be seen in **Table 2.10**.

Table 2.10. Rapid Qualitative Detection of Different Adulterants in Milk

Adulterant	**Procedure**	**Observations**	**Detection Limit (w/v)**	**References**
Urea	Take 5 mL of milk sample. Add 5 mL of p-dimethyl amino benzaldehyde reagent	Appearance of distinct yellow colour indicates the presence of added urea while a faint yellow colour indicates the natural urea in milk	0.20%	(Sharma, Rajput, & Barui, 2012)

Table 2.10 Rapid Qualitative Detection of Different Adulterants in Milk (Continued)

Adulterant	Procedure	Observations	Detection Limit (w/v)	References
Detergents	Take 5 mL of milk sample. Add 0.1 mL 0.5% Bromocresol Purple (BCP) solution	Appearance of violet colour indicates presence of detergent. Unadulterated milk will show faint violet colour	Not mentioned	(Debnath, Banerjee, Rai, & Roy, 2015)
Salicylic and Benzoic Acid	Take 5 mL of milk sample. Add 3-4 droplets of concentrated sulfuric acid. Upon acidification, add 0.5% ferric solution drop by drop and mix well	Appearance of violet colour indicates the presence of salicylic acid, while buff colour indicates the presence of benzoic acid	Not mentioned	(Debnath et al., 2015)
Formalin	Take 5 mL of milk sample. Take 1 mL of 10% ferric chloride solution in a 500 mL volumetric flask. Make up the 500 mL volumetric flask with hydrochloric acid. Take 5 mL from the mixture of ferric chloride and hydrochloric acid and put into the sample tube. Put the sample tube in boiling water bath for 3-4 mins	Appearance of pink brownish colour indicates the presence of formalin	0.10%	(Sharma et al., 2012)
Sugar	Take 5 mL of milk sample. Add 1 mL of concentrated HCl and 0.1 g resorcinol solution. Put the test tube in water bath for 5 mins.	Appearance of red colour indicates the presence of added sugar	0.20%	(Kamthania, Saxena, Saxena, & Sharma, 2014)
Starch	Take 3 mL of milk sample and mix it with 5 mL of water. Put it on boil for few mins. Cool it to room temperature. Afterward, add 2-3 droplets of iodine solution	Appearance of blue colour indicates the presence of starch	0.02%	(A. Singh, Chandra, Aggarwal, & Kumar, 2012)
Glucose	Take 1 mL of milk sample. Add 1 mL of modified Barford's reagent. Put the mixture in boiling water bath for 3 mins. Cool rapidly under tap water.	Appearance of deep blue colour indicates the presence of glucose	0.10%	(Sharma et al., 2012)
Salt	Take 5 mL of milk sample. Add 1 mL of silver nitrate solution. Mix thoroughly. Add 0.5 mL of 10% potassium chromate solution	Appearance of yellow colour indicate the presence of added salt, while red colour indicates the milk is free from added salt	0.02%	(Kamthania et al., 2014)

Despite the easiness of application for qualitative detection methods, these methods have major drawbacks. They are only valid for a limited variation of adulterants with a certain range of concentration. Additionally, these methods are not precise, and they only work if the tester knows the adulterants that they are looking for in the milk products (Reddy et al., 2017). Therefore, it is necessary to employ a more sensitive and accurate method to ensure the quality and safety of the food products. This can be largely possible with the application of quantitative detection method.

2.5.2 Quantitative Detection of Milk Adulteration

As mentioned previously in **Section 2.5**, the type of the technique applied to detect adulterants in milk products depend on the nature of adulterants in milk. Different analytical approaches have been applied for the authenticity of milk and milk products. Techniques such as capillary electrophoresis (CE), polymerase chain reaction (PCR), and enzyme linked immunosorbent assay (ELISA) are commonly used.

Capillary electrophoresis (CE) is a separation technique that provides efficient and fast separations in automated settings followed by minimum consumption of sample and reagents (García-Cañas & Cifuentes, 2008). In CE, the separation of the small and large molecules takes place in a narrow bore fused silica capillary (Cifuentes, 2006). To date, CE has been widely applied in food science as a method to detect contamination, quality control investigation, and adulteration detection. It allows rapid and reliable separation of two different proteins with high resolution and good quantification using only small amount of samples and buffers (Poonia et al., 2016).

So far, CE had been used for the adulteration detection and quantification of CM in GM (Cartoni, Coccioli, Jasionowska, & Masci, 1999), CM in SM (Trimboli, Morittu, Cicino, Palmieri, & Britti, 2017), and CM in buffalo milk (Trimboli et al., 2019). However, CE has several shortcomings. CE has a low sensitivity in analysing metabolites and impurities of a sample (Prajapati & Agrawal, 2014). Because of its low sample loading capacity, CE has ineffective interfaces between its two separation dimensions (Huang, Huang, Hu, & Chang, 2006). Other than that, CE also has a poor reproducibility (Vemireddy, Satyavathi, Siddiq, & Nagaraju, 2015). This means CE cannot produce independent and conclusive results, makingit less effective in detectingadulterations.

Polymerase chain reaction (PCR) is a DNA-based technique that is usually used to detect milk from different species as adulterants (Mohammed, 2019). It is great at detecting milk from

another species owing to its high specificity and sensitivity. In fact, PCR could detect a very low concentration of adulterants (Feligini et al., 2005). López-Calleja et al. (2005) have used PCR for the detection of GM in SM, where they found 0.1% of GM in the SM products. Bobková, Židek, Flimelová, Bobko, and Fiková (2009) employed PCR to detect the adulteration of SM with CM, where they found 8 falsified SM milk samples. In another study, PCR was also used to detect the presence of CM in GM mixture (Jung, Jhon, Kim, & Hong, 2011). Subsequently, PCR has its own limitation that preclude it from being the best method to detect adulteration in milk.

Since PCR is ideal for the detection of adulterants in natural products with intact DNA, any kind of imbalance in the DNA structure due to the presence of an inhibitor, loss, and distortion could seriously affect the result of PCR (Ambrose & Cho, 2014; Hazra, Sharma, Sharma, & Arora, 2017). Given DNA of milk-producing animals such as cow, goat, and sheep changes over time because of its environment, the result of the PCR might be affected (Liao, Liu, Ku, Liu, & Huang, 2017). Additionally, special training and equipment is needed to perform PCR. This renders PCR as an expensive and time-consuming method (B. Singh, Ganguly, & Sunwoo, 2016). Lastly, PCR can only be used for checking adulteration of milk with another milk, it cannot be used for adulteration with chemical compounds such as urea or detergents.

Enzyme-linked immunosorbent assay (ELISA) is an antibody-based analytical technique used to detect contaminants and adulterants in many food materials (Ambrose & Cho, 2014). It is one of the most common techniques used to detect foreign protein in milk authenticity. ELISA has been used extensively because of its reliability and straight-forward application (Zachar et al., 2011). Other than that, it is readily automated. Numerous studies have incorporated ELISA as a means to detect adulteration in milk. Hurley, Coleman, Ireland, and Williams (2004) had previously used ELISA as a means of detecting CM in GM, SM, and buffalo milk. Sun, Liu, Kuang, and Xu (2013) have used ELISA to analyse milk powders spiked with different levels of melamine. In another study, ELISA was used to detect the presence of CM in yak milk (Ren et al., 2014). Although it is an easy and straightforward technique, ELISA is not the best method to detect milk adulterations.

ELISA requires a labour-intensive procedure that needs a number of incubations followed by several washing steps to perform the assay (Gurevich, Kotharu, McCann, & Bertolini, 2017). Furthermore, it is also not sensitive enough in detecting low concentrations of adulterants (Ambrose & Cho, 2014; Poonia et al., 2016).

To conclude, there is a need to apply an effective and robust analytical approach to detect of adulterants in milk. This means not only focusing on the refinement of the analytical technique; it is also important to focus on the sample preparation stages which tend to be difficult to automate (Poonia et al., 2016). Automated methods will allow faster and more accurate analysis and also the avoidance of performing complex protocol of sample preparation (Das, Goswami, & Biswas, 2016). Other than that, the method will be more reliable as it has high degree of repeatability.

2.6 The Use Metabolomics in Detection of Milk Adulteration

In recent years, there has been an increased attention on the application of metabolomics to detect milk adulteration due to advancements in analytical method and computational power. Metabolomics is a fast-growing approach for the identification and quantification of small-molecule metabolites (<1500 Da) in biological samples (Consonni & Cagliani, 2010; Krishnan, Kruger, & Ratcliffe, 2004). The metabolome is defined as a pool of small molecule metabolites or chemicals found in an organ, organism, or cell (Vázquez-Fresno et al., 2014). These small molecules can include a wide variety of exogenous and endogenous chemical matters including, carbohydrate, amino acids, peptides, vitamins, nucleic acids, alkaloids, and minerals (Wishart, 2008). More information regarding the terms related to metabolomics or metabolites can be found in **Table 2.11**.

Table 2.11. Metabolomics terms and definitions

Term	Definitions
Metabolites	Small molecules that participate in general metabolic reaction required for the maintenance, growth and normal function of cells
Milk metabolome	The complete sets of metabolites in milk
Metabolomics	Identification and quantification of all metabolites in biological system
Metabolomics profiling	Quantitative analysis of set of metabolites in a selected biochemical pathway or a specific class of compound, including target analysis (limited number of metabolites)
Metabolomics fingerprinting	Unbiased method that focuses on classifying samples based on metabolite patterns or "fingerprints" of metabolites that changes in response to environmental or genetic alterations. The goal is, however, to identify discriminating metabolites

As adapted from Dettmer, Aronov, and Hammock (2007)

As mentioned previously in **Chapter 1**, the study of metabolomics could be classified into three types: untargeted, pseudo-targeted, and targeted analysis each with their advantages and disadvantages (**Figure 2.3**). The type of metabolomics technique applied in a research is usually chosen based on the objective of the study.

Untargeted approach is a comprehensive analysis focusing on the detection of as many metabolites as possible without having to identify and quantify them (Gertsman & Barshop, 2018). It is a powerful detection tool that offers the opportunity to discover unknown compounds (Dervilly-Pinel et al., 2012). The fact that untargeted metabolomics include all the compounds detected in the investigated food fractions, make it an unbiased method (Grauwet, Vervoort, Colle, Van Loey, & Hendrickx, 2014). It is also the best approach for sample comparison and discrimination analysis. Additionally, untargeted metabolomics gives large data-set output which needs to be processed through advanced analytical software. As a result, untargeted metabolomics is coupled with advanced chemometrics techniques for data analysis and interpretation (Zhang et al., 2019).

Pseudo-targeted metabolomics was just recently developed by Li et al. (2012). Based on the combination of the advantages taken from both targeted and untargeted metabolomics analysis, pseudo-targeted analysis provides a novel way for more accurate analysis (Xu et al., 2019). The application of the pseudo-targeted technique was established as a novel approach in transforming an untargeted metabolomics profiling to a pseudo-targeted method by GC-MS operating in the multiple reaction monitoring (MRM) mode. Thereafter, pseudo-targeted method has been used in other metabolomics studies and for marker selections (Chen et al., 2013; Cui et al., 2018; Tan et al., 2021; Zhou & Yin, 2016). In spite of that, the pseudo-targeted approach is novel and its ability in detecting metabolite presence in lower abundance is still lacking. The data extraction and treatment for pseudo-targeted analysis is greatly limited as there is no specific software paired to this approach (Wang et al., 2016). More information is needed before utilising pseudo-targeted analysis in the study of milk metabolomics.

The targeted methods are used broadly for the detection food contaminants as well as adulterants, by detecting and quantifying one or few known compounds in the sample (Cheah & Fang, 2020). Compared with the untargeted approach, the targeted approach is generally more complex. It requires higher level of purification followed by selective extraction of metabolites (Cevallos-Cevallos, Reyes-De-Corcuera, Etxeberria, Danyluk, & Rodrick, 2009). To put it another way, targeted metabolomics is the measurement of defined groups of chemically characterized and biochemically annotated metabolites (Roberts et al., 2012). Targeted methods could be performed through nuclear magnetic resonance (NMR), orbitrap mass spectrometry, triple quadrupole (QQQ) in the selected ion monitoring or multiple reaction monitoring modes (Cao et al., 2020). This method has resulted in higher sensitivity and better data quality, which makes targeted analysis suitable for quantifying a priori selected

metabolite. However, targeted method has low throughput. It cannot be applied when the metabolites are unknown.

Overall, the most common and generally preferred metabolomics method is the untargeted method. This is because untargeted method provides the most appropriate way to detect unexpected change in metabolites concentration (Alonso, Marsal, & Julià, 2015). Moreover, the fact that untargeted method can detect as many metabolites as possible present in a particular food matrix analysis contributes to its capability in observing unexpected changes. Thus, untargeted analysis is the best approach to detect milk adulteration.

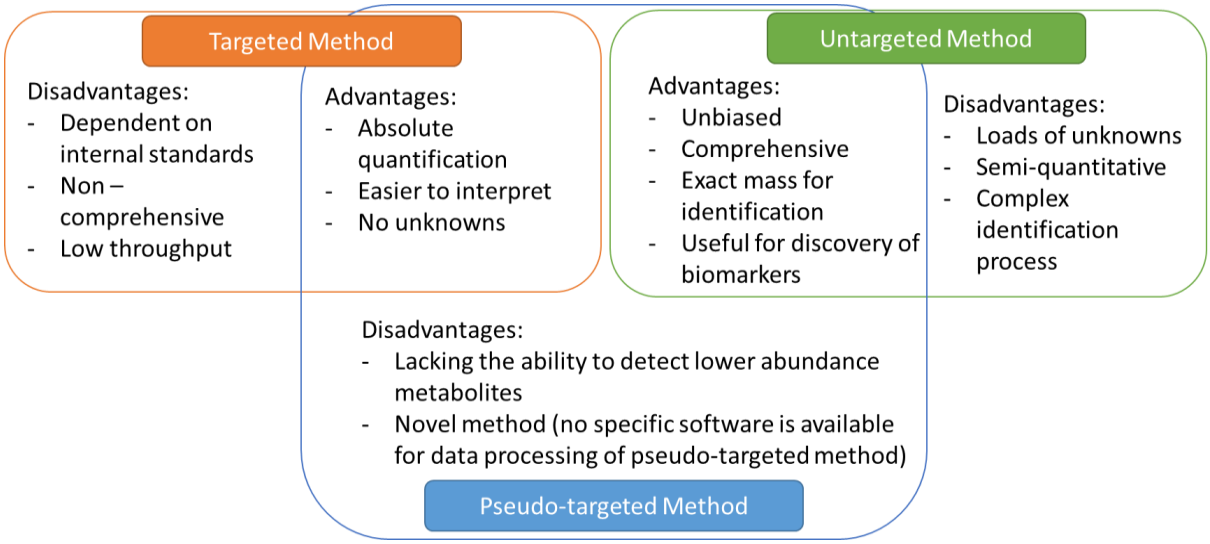

Figure 2.3. Advantages and Disadvantages of Different Metabolomics Method (targeted, pseudo-targeted, and untargeted methods)

In recent years, there has been a considerable amount of interest in the application of hyphenated techniques in milk-based metabolomics studies. This increasing amount of interest is due to improvements in the analytical performance of mass-spectrometry (MS) based method and the spectroscopy-based method (e.g., NMR). Therefore, more focus will be given to these methods in the current literature review.

2.6.1 Mass Spectrometry (MS) based milk metabolomics

Mass spectrometry (MS) is gaining increasing interest in high-throughput metabolomics and it is often coupled with other techniques such as chromatography-mass spectrometry technique (Zhang, Sun, Wang, Han, & Wang, 2012). It has a wide range of application including identification of unknown compounds, determining structure of a compound based on their fragmentation, and isotopic composition of elements in the molecule (Finehout & Lee, 2004). Due to its potential in measuring hundreds of metabolites and high sensitivity, the utilization

of the MS-based method has become increasingly popular (Koek, Jellema, van der Greef, Tas, & Hankemeier, 2011).

For the detection of milk adulteration, MS is typically coupled with chromatography and can be used for the analysis of different types of sample (i.e. liquid, gas) (Fuhrer & Zamboni, 2015). It offers numerous advantages compared with other analytical techniques including their ability in identifying different chemicals, sensitivity, and when combined with chromatography it can detect hundreds to thousands of metabolites in a given sample (Dunn, Broadhurst, Atherton, Goodacre, & Griffin, 2011). The three most common separation techniques used in MS include liquid chromatography (LC-MS), gas chromatography (GC-MS), and ultra-performance liquid chromatography (UPLC-MS)

2.6.1.1 LC-MS

Amongst the mass spectrometry methods, liquid-chromatography mass spectrometry (LC-MS) is by far the most widely used for metabolomics applications (Dettmer et al., 2007). It is highly applicable for the analysis of semi-polar compounds including secondary metabolites of interest (Zhang et al., 2012). LC provides separations of metabolites as the result of equilibration in between liquid mobile phase and a solid or liquid stationary phase. LC-MS is a sensitive and versatile method that detects both organic and inorganic molecules (Sundekilde, Larsen, et al., 2013). Additionally, the sample volume needed to run LC-MS is relatively small. About 1- 20µL of sample is enough for analysis.

The downside of LC is that it is an expensive instrumentation, and it requires an extensive and destructive sample preparation step (Finoulst, Pinkse, Van Dongen, & Verhaert, 2011). Consequently, LC methods are time-consuming, compared with GC and direct infusion or flow injection analyses (Schrimpe-Rutledge, Codreanu, Sherrod, & McLean, 2016). However, the ability of LC to increase both selectivity and data content makes it indispensable.

In milk, LC-MS was performed to identify and characterize vitamins, terpenoids, and other phytochemicals in milk (Agabriel et al., 2007). LC-MS was also used as a validation technique for milk authenticity and to determine the presence of melamine in the infant formula (Lutter et al., 2011). Additionally, Sargaeva, Wilson, and Stacey (2014) had developed method of LC-MS to measure caseins as potential biomarkers in bovine milk. Mung (2017) had developed chemical isotope labelling for LC-MS applications in the detection of adulteration in human milk. Therefore, LC-MS could also be a suitable method for the detection of milk adulteration.

2.6.1.2 GC-MS

Generally, GC-MS is used as a platform in untargeted analysis of volatile and semi-volatile compounds (Kuhara, Ohse, Inoue, & Cooper, 2011). The application of GC-MS in metabolomics analysis has proven to provide reproducible and efficient analysis. Compared with LC-MS, GC-MS achieves better metabolite separation than LC and generally avoids ion suppression, because of its use of the gaseous phase and the nature of its MS ionization. GC is more suitable for detecting volatile and semi-volatile compounds (Dunn et al., 2011). Depending on the compounds of interest, GC-MS requires extensive sample preparation steps that include chemical derivatization of the metabolic species prior analysis (Koek et al., 2011). Thus, this limits its applicability to metabolomics.

GC-MS is an excellent tool in separating, detecting and quantifying volatile compounds. Depending on the sample preparation conditions, GC-MS can be applied to the analysis of a wide range of metabolite classes including ketones, aldehydes, alcohols, esters, sulphides, sugars, sugar-phosphates, sugar-alcohols, organic acids, amino acids, lipids, peptides, alkaloids, amines and amides (Kamal & Karoui, 2015).

In milk, GC-MS had previously been used to determine volatile compounds in CM (Toso, Procida, & Stefanon, 2002), determine the presence of melamine in CM powder and infant formula (Lutter et al., 2011), characterize and compare the metabolites present in GM and GM adulterated by CM (Scano, Murgia, Pirisi, & Caboni, 2014), and compare the metabolites profile of yoghurts made from SM and GM (Murgia, Scano, Cacciabue, Dessì, & Caboni, 2019). With this, GC-MS could also be applied for milk characterization and detection of adulteration.

2.6.1.3 UPLC-MS

In metabolomics studies, UPLC-MS is a powerful approach that can be used to quantify metabolic signalling in a comprehensive manner (Taleuzzaman, Ali, Gilani, Imam, & Hafeez, 2015). UPLC works on the same principle as HPLC. The only difference is the particle size of the column material in ULC is less than 2 µm (Swartz, 2005). Therefore, compared with HPLC, UPLC-MS has higher specificity and peak capacity that makes it suitable for use in metabolomics. According to Zhang et al. (2012), the combination of UPLC and MS makes it possible to detect a number of polar metabolites and thus enlarges the number of detected analytes.

In contrast to GC-MS and LC-MS, the application of UPLC-MS is still relatively new. For metabolomic analysis, UPLC is often coupled with quadrupole-time-of-flight (Q-TOF) MS, making it possible to identify trace components in a complex mixture accurately, followed by the analysis of mass measurement that is less than 5 ppm (Su et al., 2020).

In milk, UPLC-MS was previously used to uncover the differences between traditional and commercial dairy products in Russia (Pan et al., 2018). It has also been used to detect milk metabolites originating from dairy cows with subclinical and clinical mastitis (Xi et al., 2017), and most recently, to reveal changes in milk metabolome during the lactation stage in dairy cows (Zhu, Kebede, Chen, McComb, & Frew, 2020a). Thus, UPLC-MS can also be applied in detection of adulterants in milk and milk products.

2.6.2 Spectroscopy-Based Milk Metabolomics

Other than MS-based metabolomics techniques, spectroscopy-based metabolomics technique can also be applied for milk characterization and detection of adulterants in milk. Examples of such spectroscopy-based metabolomics techniques are near infrared (NIR) spectroscopy, Fourier transform infrared (FT-IR) spectroscopy, and nuclear magnetic resonance (NMR) spectroscopy.

2.6.2.1 Near Infrared (NIR) Spectroscopy

Near infrared (NIR) spectroscopy is a well-established technique used in organic chemistry to determine the existence of particular groups in the molecule in either their solid, liquid, or gas phase (Martens et al., 2018). Due to its robustness and simplicity of the instrumentation, NIR spectroscopy has been used in the food industry for at least 40 years, particularly to control and monitor the quality and processing of food products (Ellis et al., 2012).

NIR radiation is part of the electromagnetic spectrum that covers wavelength in the range of 12,500 to 4000 cm^{-1} (Tsai & Hamblin, 2017). The broad bands originated from absorptions in the overlapping wavelength are also included in the spectra. According to Wu et al. (2011), the absorptions measured by NIR mainly correspond to vibrations involving C-H, O-H, and N-H chemical bonds.

Compared to other IR analysis, NIR can penetrate various packaging materials that are transparent to NIR light (Prieto, Pawluczyk, Dugan, & Aalhus, 2017). Moreover, NIR can penetrate much deeper into an intact food sample compared to mid-IR (MIR) with wavelength of 4000 to 400 cm^{-1}. In NIR spectra, the incident ray is typically directed to the target sample and adjusted through interaction with the sample because of absorption or transmission

scattering that is affected by either chemical and/or physical disturbances at particular wavelengths (Ambrose & Cho, 2014).

NIR has previously been employed to measure the compositional parameters of milk (Melfsen, Haeussermann, & Hartung, 2011), to detect the presence urea adulteration (Khan, Krishna, Majumder, & Gupta, 2015), and confirm the authenticity of organic milk (Liu et al., 2018). NIR, MIR, and fluorescence spectroscopy were compared to demonstrate their ability in predicting the vitamins and FA contents of CM. Based on this, NIR was found to be the one with highest prediction yield for vitamins and FA contents (Soulat et al., 2020). Based on these, NIR can also be applied in detection of adulterants in milk and milk products.

2.6.2.2 Fourier Transform Infrared (FT-IR) Spectroscopy

Fourier transform infrared (FT-IR) spectroscopy is a high-throughput and non-destructive method that can be used in a wide range analysis of samples both for research and industrial application (Poonia et al., 2016). This method is performed with a long-wave infrared radiation that records absorbance or emission of a solid, liquid, or a gas in a time domain and converts it into a frequency domain by Fourier transform algorithm (Griffiths & Haseth, 2006). When a sample is probed with an infrared beam, the functional groups of the sample will absorb the radiation and vibrate in one of the recognized ways corresponds to the biochemical species of the sample (Stuart, 2005).

FT-IR is recognized as a valuable tool for metabolomic analysis because of its holistic approach. It is capable of analysing the presence of amino acids, carbohydrates, fatty acid, fats, nucleic acid, and protein in a rapid manner with a minimum amount of sample preparation needed for each run (Ellis et al., 2012).

In milk-metabolomics analysis, FT-IR was applied to detect and quantify the adulteration of CM, GM, and SM (Nicolaou, Xu, & Goodacre, 2010), to detect adulteration in camel milk with CM (Souhassou, Bassbasi, Hirri, Kzaiber, & Oussama, 2018), and to analyse the milk composition from different dairy goat breeds (Salleh et al., 2019). More recently, FT-IR was used to measure the milk metabolomics composition in goat's mammary grand under heat stress (Salama et al., 2020). Therefore, FT-IR can also be applied in detection of adulterants in milk and milk products.

2.6.2.3 Nuclear Magnetic Resonance (NMR) based milk metabolomics

Nuclear Magnetic Resonance (NMR) is a powerful technique that can detect any molecules containing one or more atoms with non-zero magnetic moments (Krishnan et al., 2004). In

NMR, isotopes with non-zero magnetic moments (^{1}H, ^{13}C, ^{14}N, and ^{31}P) and all biological molecules have at least one or more NMR signal. Each NMR signal is characterized by their fine structure, intensity, chemical shift (frequency), which reflect the nature of the perceived nuclei (Kleckner & Foster, 2011). For these reasons, NMR spectra is information rich as it contains various information regarding the identity of the sample molecules. Moreover, NMR is useful for untargeted profiling and biomarker selection in both quantitative and qualitative detection of adulterants present in food and beverages (Wu et al., 2016).

Compared with other metabolomics analysis techniques, the main advantage of NMR is that it is suitable for compounds that are difficult to ionise or require derivatisation (Viola, Tucci, Timellini, & Fantazzini, 2006). NMR also provides highly reproducible results with coefficient of variation within the range of 1–2%. Peaks in the ^{1}H-NMR spectra can be reliably assigned to specific metabolic species, based on their chemical shifts and multiple patterns. Therefore, NMR is capable of identifying and quantifying large of metabolites in parallel from a single experiment. When an advanced high-throughput NMR methodology is used, up to 200 samples can be measured within a day with the assistance of automated liquid handlers and flow-injection probes (Lindon, 2003).

The downside of NMR is that it has low sensitivity and sometimes it does not detect low amount of compounds (Fan & Lane, 2016). Compared with MS methods, NMR is not good in detecting low-abundance semi-polar and non-polar metabolites. NMR detection also requires a larger sample size, at the minimum of 500 µL (Sundekilde, Larsen, et al., 2013). Another disadvantage of NMR is that an NMR spectrometer is expensive. Even so, the running cost of NMR is quite low, partly due to the minimal sample preparation steps employed (Kamal & Karoui, 2015).

Since the NMR method exploits magnetic properties of certain atoms with non-zero moments, different isotopes in NMR are used for different types of quantification. Amongst these, hydrogen-1 (^{1}H) and carbon-13 (^{13}C) NMR spectroscopy are the most common technique used for metabolite fingerprinting and profiling in milk. ^{1}H is commonly used since it has high sensitivity, capable of identifying extended sets of metabolites (which is interesting for an untargeted fingerprinting work). ^{13}C on the other hand is effective in profiling carbohydrates, amino acids, and fatty acids in milk (Andreotti, Trivellone, & Motta, 2006; Sacchi et al., 2018). In addition to hydrogen-1 (^{1}H) and carbon-13 (^{13}C), there are fluorine-19 (^{19}F) and phosporous-31 (^{31}P) NMR spectroscopy. Even though ^{19}F has a comparable sensitivity to ^{1}H, ^{19}F can only

detect a limited range of metabolites, mostly the fluorine-containing compounds. At the same time, ^{31}P NMR spectroscopy can be used to detect milk phospholipids (Andreotti et al., 2006). Zhu et al. (2019) utilised ^{31}P NMR to detect phospholipid in powdered infant formula.

The first application of NMR for the investigation of milk properties could be dated back to 1950s (Maher & Rochfort, 2014). Since then, NMR are used in many studies involving the identification milk fatty acids, carbohydrates, free sugars, and other small molecules metabolites (Hu, Furihata, Kato, & Tanokura, 2007). Even so, the recognition for NMR as a means to detect metabolites in milk was not until 2013 (Sundekilde, Larsen, et al., 2013).

In recent years, more NMR-based milk metabolomics studies have emerged (see ***Figure 1.1***). In the study by Monakhova, Kuballa, Leitz, Andlauer, and Lachenmeier (2012), NMR was used to validate the nutritional information in milk. NMR was also used to determine the differences between milk produced by Holstein cows and other animals (Yang et al., 2016), and for detection of adulteration and milk authenticity (Li, Yu, et al., 2017). According to the study by Tenori et al. (2018), NMR is also capable of revealing the geographical origin of CM. Most recently NMR was used to observe the changes in the milk metabolome during the lactation stage of dairy cows (Zhu et al., 2020a). Thus, given its speed, accuracy and robustness, the NMR spectroscopy method has the potential to complement or even replace the application of more traditional and time-consuming techniques used for milk characterization and authentication.

2.6.3 Summary of Milk-Metabolomics Methods

In summary, untargeted analysis is considered as the best method for adulteration detection. Untargeted metabolomics analysis is capable of detecting as many metabolites as possible without having to quantify the compounds. Moreover, when coupled with chemometrics, untargeted approach can be utilised to observe the change in the metabolite concentration and identify potential biomarkers.

Amongst the untargeted analysis method, NMR and MS are amongst the most used method for the study of milk metabolomics. When, the advantages and limitations of both methods are taken into account, NMR appears to be more feasible compared with MS (see **Table 2.12**). NMR technology provides a fast method for analysing metabolites and with little or no sample preparation steps, unlike MS analysis. Therefore, NMR-based milk metabolomics was selected as an approach in the present study for the characterization of CM, GM, and SM followed by the adulteration detection of GM and SM with different concentration of CM. A summary of

the advantages and disadvantages of NMR and MS-based method for metabolomics analysis can be found in **Table 2.1.2.**

Table 2.12. Comparison of NMR and MS-based Methods for Metabolomics Analysis

	NMR	**GC-MS**	**LC-MS**
Sample Preparation	No or little sample preparation	Extraction and chemical derivatization (depending on the compound types)	Extraction
Chromatographic Separation	No separation	High-resolution separation	Medium-resolution separation
Sensitivity	mM-µM	mM-NM	mM-pM
Speed	Fast (1-5 mins)	Slow (>30 mins)	Slow (5-9 mins)
Dynamic Range	$> 10^3$	$>10^6$	$>10^6$
Quantification Accuracy	±10%	±10%	±10%
Structural Information	High	Low	Low
Significant Advantages	• High reproducibility • Suitable for compounds difficult to ionise or require derivatization	• High precision • Ideal to resolve complex biological sample • Easy metabolite identification databases	• Soft ionization • Large mass range • Easy metabolite identification databases
Significant Disadvantages	• Overlapping signals • Some chemical class not detected • Low dynamic range	• Significant sample preparation steps • Harsh ionization • Slow analysis time • Difficult to identify novel compound	• Slow analysis time • Less robust instrumentation than NMR or GC-MS • Difficult to identify novel compound

[1]Source: (Wishart, 2008),(Abu-Aboud & Weiss, 2012), and (Lankatillake, Huynh, & Dias, 2019)

More information regarding the application of different metabolomics techniques in the milk studies can be found in **Table 2.13** shown on the next page.

Table 2.13. Summary of Milk-based Metabolomics Studies Reported in The Literature[1,2]

Investigations	Metabolites	Analytical Technique	Statistics	Reference
Coagulation properties of Danish Jersey and Danish Holstein-Friesian	Carnitine, choline, citrate, and lactose	^{13}C NMR and ^{1}H NMR	PCA	(Sundekilde, Frederiksen, Clausen, Larsen, & Bertram, 2011)
Chemical isotope labelling	Amine, phenols	LC-MS	PCA, PLS-DA	(Mung, 2017)
Profiling of sheep milk and cow milk	Melamine	^{1}H-NMR	PCA, LDA	(Lamanna, Braca, Di Paolo, & Imparato, 2011)
Holstein cows and jersey milk	Isoleucine, leucine, valine, pyruvate, lactate, succinate, capric acid, linoleic acid,	^{1}H-NMR and LC-MS	PCA, OPLS-DA	(Yang et al., 2016)
Holstein cows and buffalo milk				
Holstein cows and yak milk				
Milk quality control	Lactose	^{1}H-NMR	SIMCA, PLS-R	(Monakhova et al., 2012)
Camel, cow, and mare milk	Phospholipids (LPE, PA, EPLAS, LPC, PI)	^{31}P-NMR	ANOVA	(Garcia et al., 2012)
Cow milk and infant formula	Amino acids, carbohydrates, lipids, nucleotides, energy	^{1}H-NMR	SIMCA, PCA, PLS-DA, and OPLS-DA	(Zhao, Chen, Feng, Chen, & Cai, 2017)
Changes in milk metabolomes during lactation of dairy cows	Amino acids, carbohydrates, vitamins, nucleic acids-related compounds, and fatty acids	^{1}H-NMR and UPLC-QToF-MS	PCA, PLS-R	(Zhu et al., 2020a)
Cow milk and goat milk	Lipids (ARA, DHA, EPA)	UPLC-MS	PLS	(Li, Zhao, et al., 2017)

[1]Analytical technique: NMR (Nuclear Magnetic Resonance), UPLC (Ultra Performance Liquid Chromatography), LC (Liquid Chromatography), MS (Mass Spectrometry), GC (Gas Chromatography), QToF (Quadrupole Time of flight), NIR (Near Infrared Spectroscopy)

[2]Statistical technique: PCA (Principal component analysis), PLSDA (Partial Least Squares – Discriminant Analysis), LDA (Linear Discriminant Analysis), OPLS-DA (Orthogonal Projections to Latent Structures Discriminant Analysis), SIMCA (Soft Independent Modelling of Class Analogies), PLS-R (Partial Least Squares – Regression), ANOVA (Analysis of Variance), PLS (Partial Least Squares), CA (Cluster Analaysis)

Table 2.13. Summary of Milk-based Metabolomics Studies Reported in The Literature (continued)[1,2]

Investigations	Metabolites	Analytical Technique	Statistics	Reference
Cow milk and buffalo milk	Orotic acid, serine, threonine, valine, and urea	GC-MS	PCA, PLS-DA	(Pisano, Scano, Murgia, Cosentino, & Caboni, 2016)
Phospholipids in powdered infant formula	Phospholipids	^{31}P-NMR	PCA, PLS-DA	(Zhu et al., 2019)
Metabolites profile of goat milk and cow milk	Glycine, ribose, valine, glucose and lactose	GC-MS	PCA, OPLS-DA	(Scano et al., 2014)
Buffalo milk authenticity	Inorganic phosphate, phosphocreatine, carbohydrates, glycerophosphorylcholine, glycerophosphorylethanolamine, phosphorylcoline, and glycerol-1-phosphate	^{31}P-NMR	N/A	(Andreotti et al., 2006)
Adulteration detection in cow milk	Formalin	NIR	PCA, PLS-DA, PLS-R	(Mabood et al., 2016)
Metabolites profile and milk traits of Holstein cows	Glucopyranoside, glucosamine, ribulose-5-phosphate, phosphoenolpyruvic acid, sarcosine	GC-MS	CA	(Melzer et al., 2013)
Adulteration detection in milk	urea	NIR, Raman spectroscopy	PLS-R	(Khan et al., 2015)
Metabolomic analysis for dairy cows to measure biomarker for risk of ketosis	phosphocholine, glycherophosphocholine	1D ^{1}H-NMR 2D ^{1}H-NMR and ^{13}C-NMR	N/A	(Klein et al., 2012)

[1]Analytical technique: NMR (Nuclear Magnetic Resonance), UPLC (Ultra Performance Liquid Chromatography), LC (Liquid Chromatography), MS (Mass Spectrometry), GC (Gas Chromatography), QToF (Quadrupole Time of flight), NIR (Near Infrared Spectroscopy)

[2]Statistical technique: PCA (Principal component analysis), PLSDA (Partial Least Squares – Discriminant Analysis), LDA (Linear Discriminant Analysis), OPLS-DA (Orthogonal Projections to Latent Structures Discriminant Analysis), SIMCA (Soft Independent Modelling of Class Analogies), PLS-R (Partial Least Squares – Regression), ANOVA (Analysis of Variance), PLS (Partial Least Squares), CA (Cluster Analaysis)

2.7 Chemometrics

Since analytical techniques such as NMR and MS-based metabolomics generates complex and high-dimensional data sets, there is a need to apply an efficient and effective data analysis tools that could condense the important information and generate patterns out of the complex data (Goldrick et al., 2020). Out of many techniques, chemometrics is the most appropriate tool. Originally established in 1970s by Svante Wold, Bruce R. Kowalski, and Luc Massart, the term 'chemometrics' came from a grant application proposed by S.Wold, who thought that it was easier to receive funding for a new discipline (Héberger, 2008). Thereafter, there has been many proposed definitions regarding chemometrics. However, in the recent study by Kamal and Karoui (2015), chemometrics was defined as the branch of science that focuses on the application of mathematical and statistical methods to process data and ensure the data contain maximum information.

For detection of adulterations, chemometrics is applied when multivariate data sets are generated (Capuano, Rademaker, van den Bijgaart, & van Ruth, 2014). As the data sets generated by NMR instrumentation are huge, the data handling process has to be divided into two steps: data pre-pre-processing and data analysis (Katajamaa & Oresic, 2007). Data pre-processing involves eliminating and transforming the raw data into a new format that can be used for the data analysis step (Lawless & Heymann, 2013). Some information on chemometrics terms is found in **Table 2.14** below.

Table 2.14. Chemometrics terms and definition (As modified from Ellis et al. (2012))

Term	Definitions
Multivariate data	Data comprising of many variables collected on the samples. Often referred to as input data
Metadata	This information is used in the classification or quantification modelling. Often referred to as the output data or Y-data. In case of adulteration data, this will refer to the level of adulteration
Classification modelling	The aim of classification modelling is to classify sample into groups. In case of milk adulteration, it could be adulterated milk and unadulterated milk.
Quantitative modelling	The aim of quantitative modelling is to quantify the trait of interest. In case of milk adulteration, this could be the level or concentration of adulterant or contaminant, thus the Y-data is the level of the trait of interest (e.g., 1%, 2%,)
Unsupervised learning	Analysis performed on only the X-data with the goal of generating clusters from these input data. Often referred to as dimensionality reduction or simplification
Supervised learning	Analysis performed on both X-data and Y-data. This process involves mathematical transformation that can correlate X-data with the target trait (Y-data). It is often achieved by reducing the error between the models output prediction and the known target traits. Thus, Y-data information is essential

Based on the required information and analytical application, chemometrics can be divided into two approaches: unsupervised approach and supervised approach (Moncayo, Manzoor, Navarro-Villoslada, & Caceres, 2015). Each approach has different types of modelling method each with different aims (**Table 2.15**).

Table 2.15. Chemometrics modelling methods and their uses

Chemometrics model	**Category**	**Type**	**Uses**
Principal Component Analysis (PCA)	Unsupervised	Exploratory	PCA is used to explore the data for obvious cluster and/or detection of outliers
Hierarchical Cluster Analysis (HCA)		Summative / Exploratory	HCA is used to analyse similarities between samples. In HCA, similar samples are clustered together, while the distance between different samples are calculated
Partial Least Squares Discriminant Analysis (PLS-DA)	Supervised	Classification	PLS-DA is used to separate samples according to group classifications
Partial Least Squares Regression (PLS-R)		Quantitative	PLS-R is used to analyse and predict the dependent variable as a function of the independent variable
Artificial Neural Networks (ANNs)		Quantitative / Classification	ANNs is used to map non-linear functions from X-data to Y-data

2.7.1 Unsupervised Approach

In chemometrics, the unsupervised approach is commonly used to discover pattern in a high-dimensional dataset without using the class membership information of the sample. The unsupervised method only exploits the explanatory variable (**X**), while the supervised learning method takes both the explanatory and response variable (**Y***)* into account.

In unsupervised method, it is important to cluster and categorise the data set from the same category in order to find dissimilar samples and identify outliers (Sarker & Nahar, 2015). Two of the most widely used unsupervised approach are hierarchical cluster analysis (HCA) and principal component analysis (PCA) (see ***Table 2.15***). However, because of the the objective of the present study, more focus is given to PCA.

2.7.1.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) originated from Pearson (1901) and is one of the most commonly used chemometrics tools (Rácz, Bajusz, & Héberger, 2018). As an excellent technique for exploratory data analysis, PCA can also be used as a data mining tool. PCA can

simplify the complexity of high-dimensional data sets, while maintaining trends and patterns (Keerthi Vasan & Surendiran, 2016; Lever, Krzywinski, & Altman, 2017). Additionally, it can find patterns without prior knowledge about the reference or whether the sample have phenotypic differences, or if they came from different treatment groups.

Generally, datasets in PCA are reduced by geometrical projection analogy. Generally, the geometrical projection analogy is used to introduce derivation of bilinear data models, focusing on scores, loadings, residuals, and data rank reduction. (Esbensen & Geladi, 2009). Often called projection method, PCA looks for the direction in the multivariate space in which provide the best fit of the data distribution (Biancolillo & Marini, 2018).

In PCA, the principal components (PCs) are the linear combinations of the original variables that account for the variance in the data (Jolliffe & Cadima, 2016). Each PC is directed towards maximum variance excluding the variance already accounted for in all its preceding components (following orthogonality). Subsequently, the first component covers the maximum variance, and each component that follows it covers a lesser value of variance (Smith, 2002). The maximum number of components extracted always equals the number of variables. The PCs are interpreted based on the magnitude and direction of coefficients of the original variables. The larger the absolute value of the coefficient, the more important the corresponding variable is in calculating the component (He, Gao, Sophian, & Yang, 2017).

In standard terminology, the original matrix in PCA is decomposed into the multiplication of scores and loading matrices. The scores are interpreted as the projected samples in the new space by the new variables (i.e., PCs); while the loadings are the coefficients that is multiplied by each variable where PCs are definable as the linear combinations of the original variable (Oliveri & Forina, 2012). In PCA, it is important to know the number of PCs that must be retained to reach the maximum variation in the data (Diana & Tommasi, 2002). For this, cross-validation could be performed as it provides an estimate of the optimum number of PCs together with the expected error (Camacho & Ferrer, 2014).

After cross-validation is performed, the data visualization can be generated. The constructed PCA plots includes scores, loadings, and biplots are used as a summary about which compounds are comparable to one another and which one are different. PCA does not provide the information about what makes the samples different from each other. It cannot utilize the class labels to improve its discriminative ability (Huang, Yang, Yongxin, & Zhang, 2015). To

know more about what make the samples or compounds different from each other, the unsupervised PCA technique must be applied in combination with supervised approach.

2.7.2 Supervised Approach

In chemometrics, supervised approach is generally used to make predictions of the data output with the help of discrimination, calibration, and classification models depending on research problems (Odziomek, Rybinska, & Puzyn, 2017). In other words, it requires labelled datasets (Y-variables) to predict the output. For supervised chemometrics method, the classification is performed typically in conjunction with the use of rank reduction technique. There are few examples of supervised methods (see **Table 2.15**). However, because of the nature of the present study, more focus is given to partial least squares (PLS).

Initially suggested by Herman Wold in 1960s, PLS is derived from principal components regression that helps in building a model to predict one or more dependent variables (Peter et al., 2019). PLS algorithm is now widely applied across the field of chemometrics, neuroscience, bioinformatics, medicine, and social sciences (Tang, Peng, Bi, Shan, & Hu, 2014). Unlike PCA that derives from PCs, PLS derives latent variables (LVs). In PLS, LVs describe the maximum covariance proportion between the explanatory variable (X) and the response variable (Y) in which, the latter represent the information explained or predicted by the model (Jonsson et al., 2005). The two commonly used approaches in PLS are partial least squares regression (PLS-R) and partial least squares discriminant analysis (PLS-DA).

2.7.2.1 Partial Least Squares Regressions Analysis (PLS-R)

PLS-R is a technique that combines the features and generalizes multiple linear regression (MLR) and PCA (Guebel & Torres, 2013). The aim of this technique is to analyse and predict the dependent variable (Y) as a function of independent or predictors variables (X) (Abdi, 2010). PLS-R is capable of analysing data with many noisy and collinear variables. As a projection method, PLS-R tolerates moderate amounts of missing data in both X and Y matrices. The larger the matrices, the higher the proportion of missing data that can be tolerated (Wold, Sjöström, & Eriksson, 2001).

In PLS-R, the dependent variable (Y) can be ordinal/categorial or continuous. PLS-R is capable of modelling and analysing several Y's together, it has the advantage of giving a simpler overall picture than one separate model for each Y-variable (Carrascal, Galván, & Gordo, 2009).

The application of PLS-R is especially useful when the number of the predictor variable are highly correlated and when the number of predictors is higher than the number of the

observation variables or overfitting (Carrascal et al., 2009; Rácz et al., 2018). Since overfitting may limit the predictive ability of PLS-R, it is important to avoid overfitting. For this the step of selecting the optimum number of LVs is important (Gowen, Downey, Esquerre, & O'Donnell, 2011).

2.7.2.2 Partial Least Squares Discriminant Analysis (PLS-DA)

In the past two decades, the application of PLS-DA has become increasingly popular in the field of chemometrics, recommended for metabolomics and other 'omics' analysis (Ruiz-Perez & Narasimhan, 2018). Often called the supervised version of PCA, PLS-DA combines dimensionality reduction and discriminant analysis into one algorithm. Even so PLS-DA and PCA have different standard linear combinations (SLCs) (Maitra & Yan, 2008). In PCA, the SLCs only capture the characteristic of the predictive variables (X vector). On the other hand, PLS-DA used the SLCs to gain a substantial amount of information of the predictive variables along with the relationship between the predictive and target variables (X and Y vector) (Worley & Powers, 2013).

In theory, PLS-DA is a method of linear classification that applies the logic of PLS-R to differentiate between group of samples and classify them, in PLS-DA, the Y-variable is categorical (Rácz et al., 2018). Different from PCA, which is modelled through PCs, the variability of the data sources of PLS-DA is modelled by LVs. In PLS-DA, the LVs is the linear combination from the original variables that is used for graphical visualization of the data set (scores plot) (Pomerantsev, 2008). PLS-DA is flexible; it does not assume the data to fit a particular distribution, unlike other discriminant techniques such as Fischer's linear discriminant analysis (LDA) (Lee, Liong, & Jemain, 2018).

As a classifier and feature selector method, PLS-DA has few advantages. It has the ability to handle the noisy collinear variables that are often found in the 'omics' datasets (Lämmerhofer & Weckwerth, 2013). Furthermore, the result of the data is typically summarised by PLS-DA in the form of scores and loading plots (Rodríguez-Pérez, Fernández, & Marco, 2018).

Nevertheless, PLS-DA has major drawbacks. PLS-DA is prone to overfitting and it could lead to false discoveries. As a solution, cross-validation needs to be performed to ensure the data are transformed to a lower dimensional space with as small error as possible (Westerhuis et al., 2008). Overall, PLS-DA is powerful in the sense that it can be utilized for both predictive and descriptive modelling and variable selections. The method is applied in the present study to classify the metabolites present in the milk products from cow, goat, and sheep.

2.8 Conclusion of Literature Review

Based on the literature review, it is known that NZ is the 8th largest milk producer in the world and the world's largest dairy exporter (DCANZ, 2020). Milk is a staple food that is rich in nutrients and it is nature's most complete food. In NZ, the milk sources are not limited to cows, milk from goat and sheep also popular.

As high value goods, NZ dairy products are susceptible as targets of adulteration and counterfeits. Since adulteration is done mostly by the addition of unknown compounds, there is a need to implement an effective and robust analytical approach to detect milk adulteration. The method must be rapid, simple, and have high reproducibility. For this, NMR-based metabolomics method has a huge potential.

Because of the large, complex, and high dimensional data generated from NMR, an advanced chemometrics methods must be applied to allow comprehension and understanding of the data. An advanced chemometrics algorithm had shown a great ability in unfolding the high dimensional data and select interesting biomarkers (Cui, Zhang, Cai, & Shao, 2017).

Chemometrics is classified into two different approaches: unsupervised and supervised. Unsupervised data analysis is usually implemented to find the hidden data and select interesting biomarkers. In the unsupervised approach, there are no output variables to predict; instead the objective is to find patterns present in the data based on the relationship of the data with each other (Kotu & Deshpande, 2019). On the other hand, supervised data analysis is used to predict the value of the output variables based on the series of input variables. In order to perform this, the model is developed from the training set where the values of both input and output variables are previously known (Moncayo et al., 2015).

2.9 Research Gaps

Gathering all the information about the compositional information about milk and milk metabolomics study from the literature study, three research gaps were identified.

1. Despite the popularity of GM and SM, most of the study conducted around milk metabolites are heavily focused on CM. As a result, there is not much information regarding the metabolite properties of GM and SM. For this reason, there is a need to conduct more studies regarding GM and SM.

2. Although NZ plays a key role in the world's milk industry, there is not enough information regarding the characteristic of NZ CM, GM, and SM. Thus, a study to characterize the metabolite profile of NZ CM, GM, and SM is needed.

3. Despite the effectiveness of ^{1}H-NMR in metabolomics analysis, up to date there have only been 23 publications (found in Webs of Science) on the application of ^{1}H-NMR in milk-based metabolomics studies (see ***Figure 1.1***). As for the application of ^{1}H-NMR in detection of milk adulteration, there were only seven publications. As a result, the application of ^{1}H-NMR in milk metabolomics, particularly in detection of adulteration is still relatively novel and more studies are needed.

Chapter 3 . Objectives, Research strategy and Overall Experimental Approach

3.1 Objectives of the study

The general objective of the study is to find out whether NMR-based metabolomics methods combined with chemometrics is suitable to characterise and detect adulteration in NZ GM and SM with different concentrations of CM.

The specific objectives in the study are as follows:

1. To characterise powdered milk from cow, goat, and sheep by the application of ^{1}H-NMR
2. To apply advanced chemometrics and feature selection to identify metabolites (biomarkers) that discriminate the different milk types from each other
3. To detect adulteration in GM and SM through the application of metabolomics method

3.2 Research Strategy

The present study was done based on the application of the emerging field of metabolomics with advanced chemometrics, with a focus on proton Nuclear Magnetic Resonance (^{1}H-NMR) spectroscopy. Together, metabolomics and chemometrics are excellent tools for characterization and detection of new food ingredients. Hence, metabolomics and chemometrics are employed in this study to characterize New Zealand cow, goat, and sheep milk and select potential biomarkers for detection of adulterations.

For data extraction and interpretation, unsupervised (PCA) and supervised chemometric approaches (PLS-DA and PLS-R) were used.

3.3 Experimental Approach

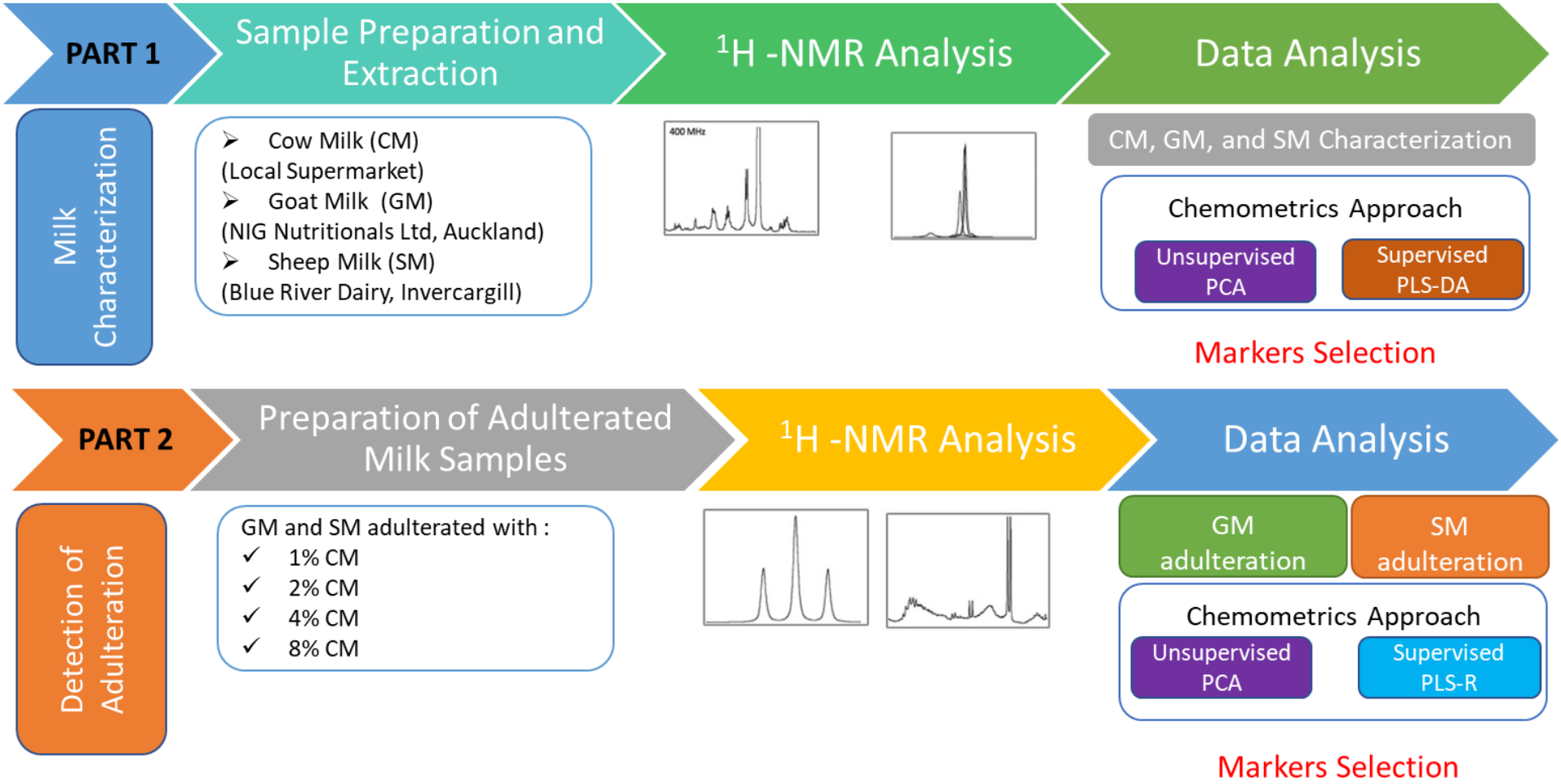

Figure 3.1. Experimental approach for characterization and detection of adulterations of New Zealand cow, goat, and sheep milk

Chapter 4 . Characterisation and Identification of New Zealand Milk (Cow Milk, Goat Milk, and Sheep Milk)

4.1 Introduction

As the world's largest exporter of dairy products, New Zealand (NZ) exports about 95% of its dairy products (including milk, butter, cheese, and whole milk powder) to more than 100 countries (McGiven, 2016; NZIER, 2017). Some of the destinations of these dairy products includes countries with questionable food safety and quality controls standards (e.g., China, India, Iran and Sudan) where there have been numerous reported incidents of milk adulteration (Moosavy, Kordasht, Khatibi, & Sohrabi, 2019; Salih & Yang, 2017). In most cases, the adulteration of milk and its products is economically motivated. Economically motivated adulteration (EMA) is defined as the fraudulent and intentional addition or substitution of a substance in a product for the purpose of increasing the product value while reducing the production costs for economic gain (Everstine et al., 2013; Moore et al., 2012). Other possible reasons for adulteration of milk include to conceal the perishability and extend the shelf life of milk, to fulfil the demand for milk products that was not met, and to replace a premium source of milk (e.g., goat and sheep milk) with cheaper source of milk (e.g., cow and soy milk) to gain more profit.

Some common parameters used to evaluate milk quality are protein content, freezing point, fat content, and solid-non-fat (SNF) percentage. For these reasons, several common adulterants are added into milk to manipulate the milk quality results so properties such as nutrient content can appear higher than they really are (Azad & Ahmed, 2016). For example, melamine is added to artificially increase the protein content in milk, while formalin and caustic soda are sometimes added to extend the shelf life of milk. Similarly, urea is added to increase SNF and NPN contents (Gabriels, Lambert, Smith, Wiesner, & Hiss, 2015). In some cases, cheaper milk from one species is also used as adulterants in milk from another species to minimise cost and maximise the profits. Although adulterants are added in milk to increase its value, most of them pose potential health risks (see **Section 2.4.1**).

When milk is adulterated, not only does it become inferior in quality, adulteration of NZ milk will also have a financial impact on the business, followed by damage to the brand reputation and potentially legal consequenceswhen the product causes adverse health effects. The products will have to be recalled and the manufacturer will lose the distributor's and the public's trust (Handford et al., 2016). In a worst case, it could also cause bankruptcy to the

company like what happened to San Lu. In 2008, the infant milk powder produced by San Lu was adulterated with melamine causing the hospitalisation of 300,000 babies and the death of many (Xin & Stone, 2008). Given their established 'clean-and-green' reputation, NZ milk products are at a risk of adulteration. Also, NZ produces high-value milk from cow, goat, and sheep. This means, strategies for controlling adulteration are very important for the NZ dairy industry.

To prevent milk adulteration from happening, the first step is to understand the compounds that are present in the milk and their compositional properties, followed by the development and validation of method for milk adulteration detection. To achieve these, a robust method with low recurrent costs and high reproducible results is needed. One of such method is metabolomics. In the recent years, metabolomics has been applied extensively for the detection of food frauds, with NMR being the most commonly used approach (Sargaeva et al., 2014; Sundekilde, Larsen, et al., 2013). NMR is a powerful tool that is extremely rich in information, and it can provide detailed information regarding the dynamics, structure, reaction state, and a chemical environment of a compound (García-Cañas, Simo, Castro-Puyana, & Cifuentes, 2014). Because of this reason, advanced chemometrics tools including unsupervised and supervised method needs to be applied to analyse the data (S. Brown, Tauler, & Walczak, 2020). Unsupervised methods (e.g., PCA) are used to simplify the data and find a pattern without prior knowledge of the sample, while supervised methods (e.g., PLS-R and PLS-DA) are used to make predictions of the data output with the help of discrimination, calibration, and classification models while taking the sample information into consideration (Godoy, Vega, & Marchetti, 2014).

To date, not many studies are found around the characterization of NZ milk powder (cow and sheep) nor the application of NMR-based milk metabolomics to detect metabolites in NZ SM. Additionally, there is no previous study in the literature comparing the metabolites properties amongst CM, GM, and SM. The previous studies conducted from the same research group were about the application of NMR in detecting the metabolites present in NZ GM (Sanchez, Zhu, Frew, & Kebede, 2020) and the changes in milk metabolome (CM) during lactation stage of dairy cows (Zhu et al., 2020a). Sanchez et al. (2020) had identified the presence of 44 metabolites in NZ GM powder, including high levels of citric acid, valine, pantothenate, creatinine, and valine. On the other hand, Zhu et al. (2020a) identified 18 metabolites including organic acids, vitamins, carbohydrates, fatty acid, and amino acids and derivatives from the CM metabolome.

Therefore, in the current study, NMR is combined with advanced chemometrics to identify the metabolites present in milk powders coming from NZ CM, GM, and SM. The first step was to characterise different types of milk powders, according to their metabolite composition. After the milk metabolites were successfully identified, potential biomarkers that makes the milks different from each other were selected using advanced chemometrics and variable identification methods. In this study, the biomarkers were selected based on the calculated variable identification (VID) coefficients. Therefore, the current study is the first one to compare the metabolites present in NZ CM, GM, and SM by the application of ^{1}H-NMR.

4.2 Materials and Method for Milk Characterization

The materials and method used in the study are described as follows.

4.2.1 Samples

Three different kinds of milk powder from cow, goat, and sheep were obtained from different manufacturers in New Zealand. The CM powder was obtained from a local supermarket in Dunedin, NZ. The GM powder was a kind donation from NIG Nutritionals Limited, Auckland, NZ. The SM powder was a donation from Blue River Dairy, Invercargill, NZ. Upon arrival, these samples were subdivided and transferred into 50 mL falcon tubes. A preliminary study on freeze-dried milk was done by Zhu, Kebede, Chen, McComb, and Frew (2020b), where they found that samples stored at -20ºC showed excellent metabolites stability for at least 224 days. Therefore, the samples in this study were stored in the freezer at -20ºC until analysis. Before analysis, these frozen samples were thawed for at least an hour in a cooling room at 4ºC. All analysis were done using the same batch of milk powder.

4.2.2 Reagents

Acetonitrile (ACN) was obtained from Fischer Scientific (Waltham, MA, USA). The organic solvents, methanol, ethanol, and chloroform (HPLC grade) were obtained from Thomas Scientific (Swedesboro, NJ, USA). Deuterium oxide (D_2O; D, 99.9%) was obtained from Cambridge Isotope Laboratories (College Park, MD, USA). Phosphate buffer (PB) and sodium 3-(trimethylsilyl) propionate-2,2,3,3-d4 (TSP) were purchased from Sigma-Aldrich (St Louis, MO, USA). Lastly, TSP was dissolved in D_2O at a concentration of 3 mmol L^{-1}.

4.2.3 Sample Extraction and Preparations for ^{1}H-NMR analysis

Sample extraction is one of the critical steps in metabolome analysis. The method used for sample extraction and preparations was based on the study by Zhu et al. (2020b) with several modifications. To eliminate high molecular weight compounds that may disguise the signal of

the metabolites in different milk samples (lipid fractions), different organic solvents were used to extract compounds form the powdered milk sample. Different types of solvents including methanol, chloroform, methanol-chloroform mixtures, and ethanol were tested for their capabilities in extracting high numbers of metabolites presents in the freeze-dried milk samples. Amongst all the solvents, methanol was found to be the best, as it produces high-quality spectra with good peaks visibility. Similarly, Sanchez et al. (2020) had performed method optimization for ^{1}H-NMR analysis where methanol was found to be the best extraction solvent. Thus, methanol was chosen as the final extraction solvent for the current ^{1}H-NMR analysis.

For sample preparations, 1 g of the milk powder was dissolved in 4 mL of Milli-Q water and centrifuged at 8,225 x g for 30 min at 4ºC (Heraeus centrifuge 16R, Thermo Fisher Scientific, UK). The sample was centrifuged for the separation of fat and other impurities. An aliquot (0.5 mL) of the supernatant was then transferred into a new centrifuge tubes, followed by the addition of 1 mL methanol. The mixture was then vortexed for 10 minutes and centrifuged again at 8,225 x g for 10 min at 4ºC. Afterwards, the supernatant was taken and dried under a stream of nitrogen.

The dried samples were redissolved in 700 µL D_2O (0.1 M phosphate buffer (PB), pH=7.4, 5 mM TSP). The PB was used to maintain the pH of the milk extracts to prevent signal shifts, while TSP in the buffer acted as the reference for the calibration of NMR shift. After a further centrifugation at 8,225 x g for 5 min at 4 °C, 600 µL of the sample was transferred to a 5 mm NMR tube (5 mm; Norell ST500; Norell Inc., Morganton, NC, USA) for analysis. All the sample analysis was performed in quadruplicates.

4.2.4 1H-NMR Experiments

The NMR experiments were performed on a Varian spectrophotometer at a resonance frequency of 400 Mhz. The spectra for ^{1}H NMR were acquired with 90° pulse sequence, 2 s relaxation delay and 128 scan numbers, requiring about 12 minutes per sample. The NMR procedure had previously been optimized by Zhu et al. (2020b) and Sanchez et al. (2020).

4.2.5 1H-NMR Data Pre-processing

Prior to multivariate analysis, the NMR data was pre-processed for the removal of unwanted variation such as instrumental or experimental artefacts. The data pre-processing was carried out using the commercial software suite MestReNova (version 12.0.3, 2018, Mestrelab Research, Santiago de Compostella, Spain). To correct distortions between the peaks, manual

baseline correction and auto phase correction were performed. Afterwards, the spectra were then referenced to TSP at the chemical shift (δ) of 0.0 ppm.

4.2.6 ^{1}H-NMR Analysis

Identification of impurities and solvent peaks were done, and those peaks were disregarded prior to the identification of possible metabolites present in the NZ CM, GM and SM samples with multiplet analysis. Next, the characterization of metabolites present in the NZ CM, GM, and SM samples were performed by comparing the chemical shifts of the present NMR spectra with the data from the past studies and online databases such as the Human Metabolome Database (HMDB), the Milk Composition Database (MCDB), and the Chenomx NMR Suite 8.43.

4.3 Multivariate Data Analysis

Both unsupervised PCA (see **Section 2.7.1.1**) and supervised PLS-DA (see **Section 2.7.2.2**) multivariate data analysis (MVDA) were employed to obtain a model to effectively classify the metabolites present in NZ CM, GM and SM. Accordingly, VID coefficients were calculated in order to select discriminant compounds that can be used as potential markers of NZ CM, GM and SM.

4.3.1 Unsupervised Principal Component Analysis (PCA)

PCA is an excellent technique for data exploration and pattern recognition. It allows analysis of datasets that might contain imprecise measurements, categorical data, and even missing value as PCA does not use class label information (Worley & Powers, 2013). Consequently, the NMR data set obtained were used as inputs for the data matrix **X**. A software program called SOLO (Version 8.6, 2018, Eigenvector Research, Wenatchee, WA, USA) was used for data processing and visualisation. In the SOLO software, a 'leave one out' cross validation was applied, and cumulative variance graph and the root mean square error of cross validation graph were constructed to select the optimum number of principal components (PCs). Based on the chosen PCs, scores, loading, and bi-plot were generated.

4.3.2 Supervised Partial Least Squares-Discriminant Analysis (PLS-DA)

PLS-DA is a versatile algorithm used for descriptive and predictive modelling, in addition to discriminative variable selection (Lee et al., 2018). In the present study, the NMR data set was used for both X and Y inputs in the SOLO software (Version 8.6, 2018, Eigenvector Research, Wenatchee, WA, USA). A 'leave one out' cross validation was applied to obtain both cumulative variance and root mean square error cross validation graph. Unlike PCA that uses

PCs, PLS-DA uses LVs. LVs (latent variables) in PLS-DA is used to generate scores, loadings, and bi-plot. The optimum number of LVs explains the highest variance at the lowest root mean square error based on the cross-validation graph.

4.3.3 Discriminant Markers Selection

To select the metabolites compounds that can classify the three different milk types from one another, variable identification (VID) coefficients were calculated. These values correspond to the correlation coefficient between each original X-variables (i.e., metabolites) and Y-variables (milk samples) as predicted by PLS-DA model (Kebede et al., 2014; Ooms, 1996). In this current study, only compounds with VID value higher than 0.700 were selected. For each discriminant markers, a box and whisker plot were constructed individually using RStudio Desktop (Version 1.3.1093, RStudio Team, Boston MA, USA). Each box plot displayed the five-number summary of the relative number of compounds present in the three milk types (CM, GM, and SM). The five-number summary includes the minimum, first quartile, median, third quartile, and the maximum relative amount of the selected discriminant compounds. For further analysis, SPSS Software Version 24 (IBM Corporation, New York, New York, United States) was employed to perform Tukey's range test. The Tukey's range test was used to see whether the compound has a significant presence ($p < 0.05$) among NZ CM, GM, and SM samples.

4.4 Results and Discussion

The results and discussion sections for the characterisation and identification of CM, GM, and SM are divided into four different sub-sections as follows:

4.4.1 Identification and characterization of NZ cow milk (CM), goat milk (GM), and sheep milk (SM) liquid fraction with ^{1}H-NMR

In the present study, the representative NMR spectra of NZ CM, GM, and SM are shown in **Figure 4.1**. The spectra were carefully inspected to ensure that the information of the NMR chemical shifts was accurate given the complexity of the identification process.

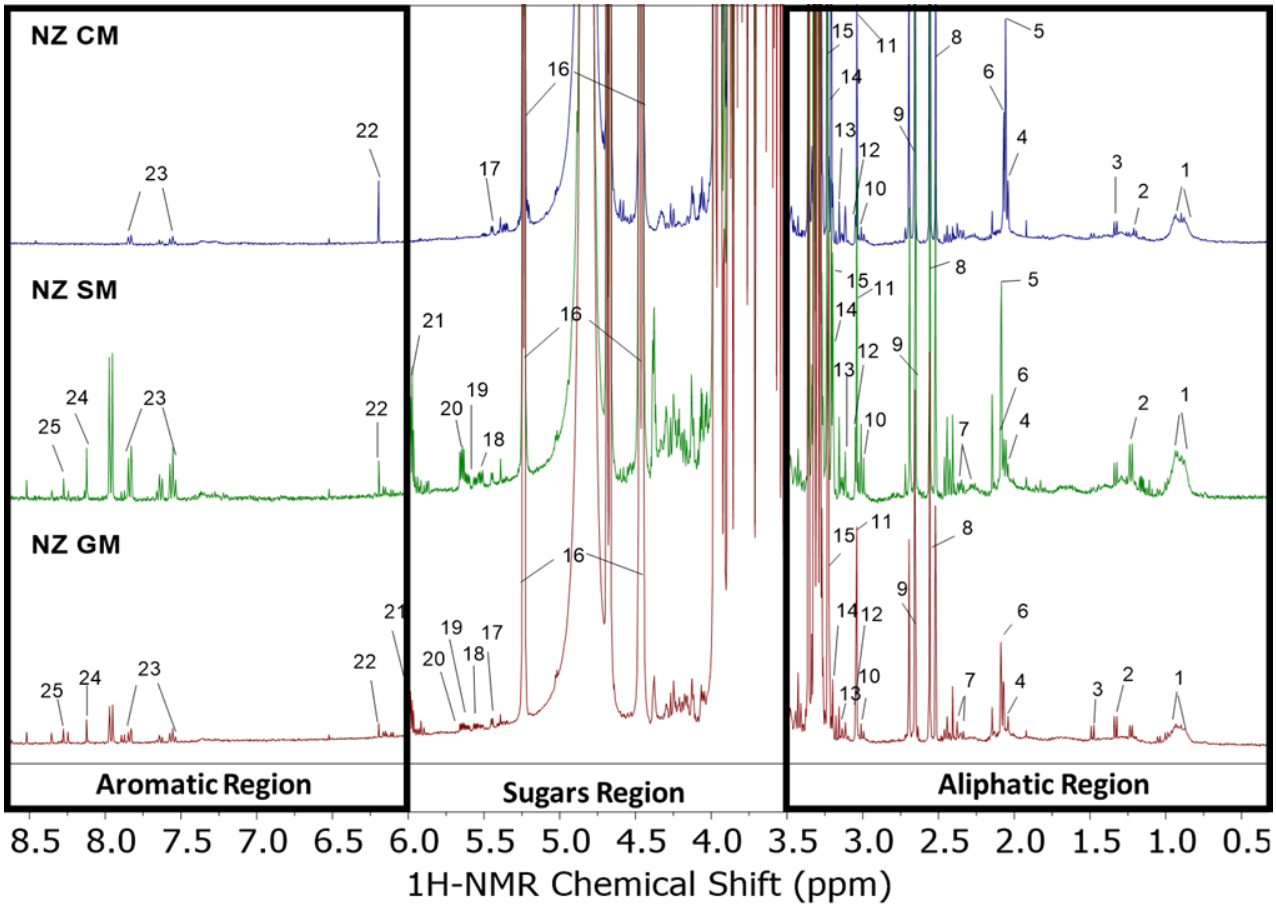

Figure 4.1. **^{1}H**-NMR Spectra of NZ CM, GM, and SM (removed lipid fractions)

As seen in **Figure 4.1**, the spectra are divided into three main spectral regions: aliphatic region, sugars region, and aromatic region. According to Dona et al. (2016), it is important to identify different regions in the spectra as it eases the process of turning classification of signals into identification of metabolites. The first region, (the aliphatic region), mainly consists of amino acids, carboxylic acids, and carbohydrates that ranges from 0 ppm to 3.5 ppm. Next is the sugar region that consist of both simple and complex carbohydrates, and ranges from 3.5 ppm to 6.0 ppm. Lastly, the aromatic region consists of carboxylic acids, nucleobases, and nucleotides and ranges from 6.0 ppm to 8.5 ppm.

A total of 25 metabolites were identified in the current study of NZ CM, GM, and SM. The small numbers of the identified metabolites found in this study is because the lipid fractions of the milk samples (CM, GM, and SM) were not considered. As a result, no fatty acids were found in the samples. The removal of the lipid fraction of the milk samples was done since lipid fractions contain compound with high molecular weights and these could potentially mask lower-abundance metabolites of interest in other areas of the spectra (Agudiez et al., 2020). Additionally, metabolomics analysis of the water-soluble fractions has been effectively used

for the authentication and adulteration detection studies of dairy products (Nascimento et al., 2017).

Another finding from the spectra is that the number of metabolites present in each spectra region were different depending on the milk type. NZ CM was seen to have 14 metabolites in the aliphatic region, followed by 2 in sugar region, and another 2 in the aromatic region. On the other hand, NZ GM was seen to have 14 metabolites in the aliphatic region, followed by 6 metabolites in the sugar region, and 4 in the aromatic region; while NZ SM have 14 metabolites in the aliphatic region, followed by 5 metabolites in the sugar region, and 4 in the aromatic regions. This initial visual investigation indicates possible differences among the metabolites of the three milk types.

Moreover, strong intensity signals were predominantly found in both aliphatic and sugar regions in all three milk types (CM, GM, and SM). These corresponds to succinic acid, citrate, phosphocholine, carnitine, and lactose. This observation is supported by the findings by Klein et al. (2010), who reported that CM has high concentrations of phosphocholine, carnitine, and lactose. In another study, Sanchez et al. (2020) showed that GM powder spectra were dominated by the signals coming from aliphatic and sugar regions, including carnitine and lactose. About SM, the previous study regarding the metabolomics comparison of SM and GM by Caboni et al. (2016) were done by GC-MS, not NMR. Therefore, this study is the first one to compare the metabolites present in CM, GM, and SM by the application of ^{1}H-NMR.

In the past studies on NMR-milk metabolomics, only identification and quantification of high-intensity peaks were done; low-intensity peaks were not considered (Li, Yu, et al., 2017; Yanibada et al., 2018). According to Dona et al. (2016), the identification of metabolites that are present at a relatively low level is deemed to be difficult since the signals could be either partially or completely overlapped and great care is required during direct integration and quantification of the spectrum. Even so, Hu, Furihata, Ito-Ishida, Kaminogawa, and Tanokura (2004) claimed that weak signals coming from NMR spectra could be caused by the large molecular weight of the proteins, leading to a lower molar concentration of the compounds represented by the weak signals. But for the purpose of adulteration detection, low intensity peaks are also important as all metabolites profile are useful as predictive biomarkers in the dairy industry (Caboni et al., 2016). For example, Klein et al. (2012) found that the presence of metabolites such as β-hydroxybutyrate, glycerophosphocholine, and phosphocholine could be used as an indicator of ketosis in dairy animals. In another study, high concentrations of β-

hydroxybutyrate, lactate, and acetate are associated with high levels of somatic cells count (SCC) in milk, while hippuric acid and fumarate are found in a lower concentration with high levels of SCC (Sundekilde et al., 2013). High SCC is an indicator of mastitis infections in cows, goats, and sheep. Hence, it is crucial to identify as many metabolites (including ones with the low-intensity signals) as possible in the CM, GM, and SM samples. The detailed information on the metabolites presents in the different milk samples including their chemical shifts, multiplicity, assignments, and compound classes can be found in **Table 4.1**.

Table 4.1. Metabolites assignment from [1]H-NMR spectra of water-soluble fraction from NZ CM, GM, and SM

Peak No	Identity	δ_{1H} (multiplicity)[1]	Assignment	Compounds Classes	Presence	Ref[2]
1	α-hydroxybutyrate	0.88(t)	CH_3	Hydroxybutyric acid	All Milk	[1]
2	Lactate	1.34(d)	CH_3	Carboxylic acid	All Milk	[1],[5]
3	Alanine	1.47(d)	CH_3	Amino acid	CM, GM	[2],[4]
4	N-acetyl glucosamine	2.04(s)	CH_3	Carbohydrate	All Milk	[6],[7]
5	N-acetyl carbohydrate1	2.06(s)	CH_3	Carbohydrate	CM, SM	[3]
6	N-acetyl carbohydrate2	2.07(s)	CH_3	Carbohydrate	All Milk	[6]
7	Glutamate	2.36(m)	CH_2	Amino acid	GM, SM	[3]
8	Succinic acid	2.56(s)	CH_3	Carboxylic acid	All Milk	[1]
9	Citrate	2.65(d)	CH_2	Tricarboxylic acid	All Milk	[4]
10	Creatine	3.03(s)	CH_3	Amino Acid	All Milk	[2]
11	Phosphocreatine	3.04(s)	CH_3	Amino acid	All Milk	[1],[6]
12	Creatinine	3.05(s)	CH_3	Amino acid	All Milk	[5]
13	Malonic acid	3.12(s)	CH_2	Dicarboxylic acid	All Milk	[2],[3]
14	Phosphocholine	3.20(m)	CH_3	Amino acid	All Milk	[1]
15	Carnitine	3.23(s)	CH_3	Amino Acid	All Milk	[7]
16	Lactose	4.45(d)	CH	Carbohydrate	All Milk	[1],[3]
		5.23(d)	CH	Carbohydrate	All Milk	[1],[4]
17	Galactose-1-phosphate	5.45(dd)	CH	Carbohydrate	CM, GM	[1]
18	Glucose-1-phosphate	5.56(dd)	CH	Carbohydrate	GM, SM	[6],[7]
19	UDP glucoronate	5.61(dd)	CH	Carbohydrate	GM, SM	[2]
20	UDP galactose	5.65(dd)	CH	Carbohydrate	GM, SM	[1],[7]
21	Guanosine monophosphate	5.97(d)	CH	Nucleotide	GM, SM	[1],[7]
22	Orotate	6.20(s)	CH_3	Carboxylic acid	All Milk	[2],[3],[6]
23	Hippuric acid	7.57(m)	CH_2	Carboxylic acid	All Milk	[3],[5]
		7.83(dd)	CH_2	Carboxylic acid	All Milk	[1]
24	Adenine	8.12(s)	CH	Nucleobases	GM, SM	[2],[3]
25	Inosine	8.3(s)	CH	Nucleosides	GM, SM	[5],[6]

[1]Multiplicity: s=singlet; d=doublet; t=triplet; dd=double of doublets; m=multiplets. δ1H (multiplicity) for the unknown signals (U1-U5) were given as follows: U1=5.39(s), U2=5.51(d), U3=5.99(t), U4=7.96(d), and U5=8.28(s).

[2]References: [1]=Human Metabolome Database (HMDB; http://hmdb.ca/); [2]=(Klein et al., 2010); [3]=(Sundekilde, Larsen, et al., 2013), [4]= (Sundekilde, Poulsen, Larsen, & Bertram, 2013), [5]=(Li, Yu, et al., 2017),[6]=(Zhao et al., 2017),[7]=(Sanchez et al., 2020)

Based on **Figure 4.1**, the metabolites found in the ^{1}H-NMR spectra of NZ CM, GM, and SM were categorised into different classes of compounds. These compounds include amino acids, carbohydrates, carboxylic acids, dicarboxylic acids, nucleobases, nucleosides, nucleotides, and tricarboxylic acids. A summary of the compounds found in the different milk types (CM, GM, and SM) can be seen in the Venn diagram (**Figure 4.2**).

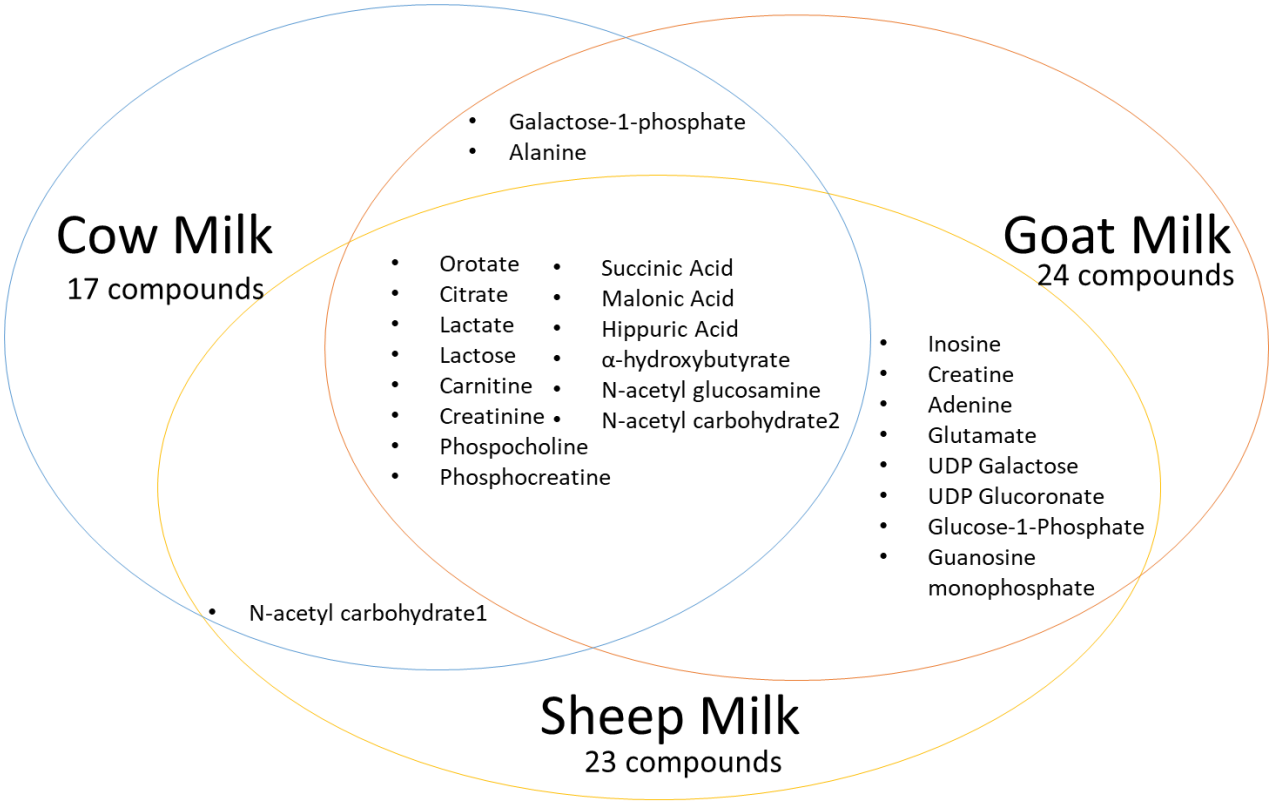

Figure 4.2. Venn Diagram for metabolites compounds found in CM, GM and SM

To provide more comprehensive information on metabolites that are present in different milk samples, advanced chemometrics (unsupervised PCA and supervised PLS-DA) was applied. Further analysis was also employed to determine potential (bio)markers for each milk type by the calculation of VID coefficient, followed by the explanation regarding the importance of the selected milk metabolites as discriminant markers. The obtained result coming from these methods are explained systematically in the following sections.

4.4.2 Unsupervised PCA Analysis of NZ CM, GM, and SM

In the current study, PCA transformed the original measured variables (i.e., NMR metabolites) into the new orthogonal or uncorrelated variable also known as principal components (PCs). Each PC is a linear combination of the measured NMR metabolites. PC 1 represent the maximum total variance. PC 2 is orthogonal to PC1, in which it represents the maximum

residual variance until the total variance is accounted for (Berrueta, Alonso-Salces, & Héberger, 2007). In other words, it is important to choose the optimum number of PCs in PCA.

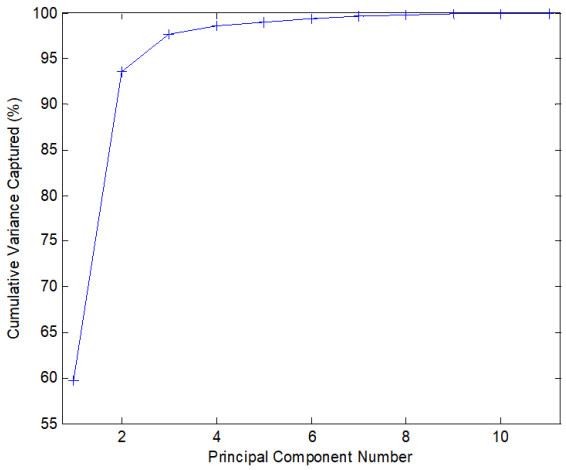

Figure 4.3. Cumulative Variance Graph - NZ CM, GM, and SM Characterization (PCA)

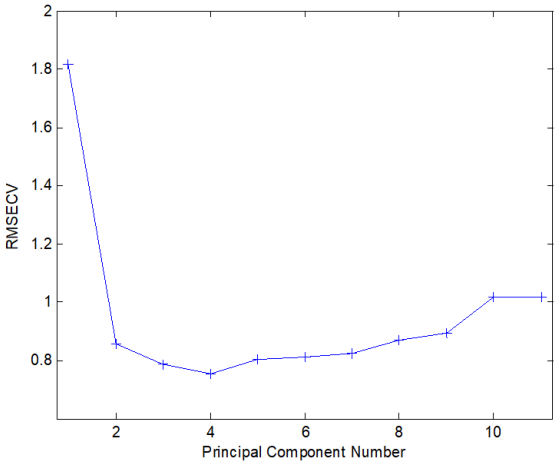

Figure 4.4. Root Mean Square Error of Cross Validation (RMSECV) Graph – NZ CM, GM, and SM Characterization (PCA)

In the present study, a leave-one out cross validation (LOOCV) was used to obtain cumulative variance and root mean square error values (see **Figure 4.3** and **Figure 4.4**). LOOCV is a special case of cross validation where the number of folds is equal to the number of instances in the dataset (Sammut & Webb, 2010). Based on the cross-validation results (**Figure 4.3** and **Figure 4.4**), 4 PCs were chosen. The 4PCs were chosen as it retains the maximum amount of information followed by minimal error or noise; aiming to keep the risk of overfitting to minimum. In this case, 4PCs explained 98.57% (PC1=59.64, PC2=33.91, PC3=4.08, and PC4= 0.93) cumulative variance of the data. Based on the chosen PCs, a score plot was generated (**Figure 4.5**).

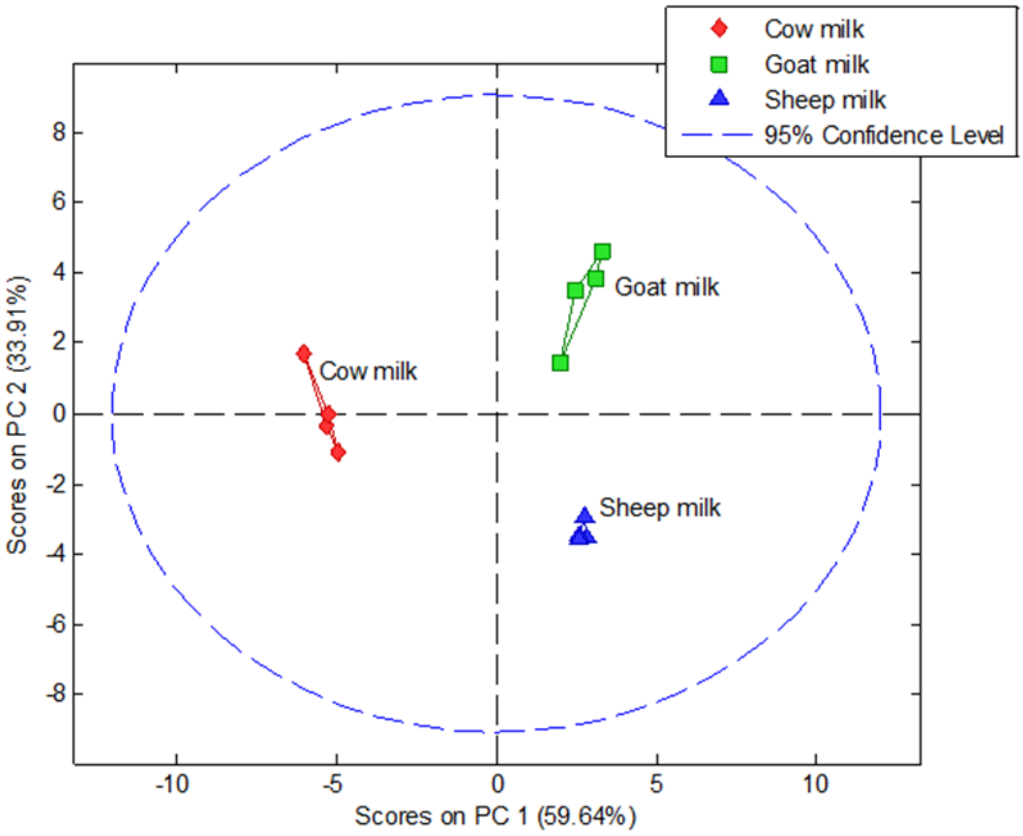

Figure 4.5. A PCA Score Plot on the Characterization of NZ CM, GM, and SM

The PCA score plot is constructed when the data matrix is decomposed into *n* rows (samples) and *p* columns. In this study, the score plot is used to demonstrate the relationship among the three milk types. Based on the distance among the milk types, one can clearly see that all the three milk types are different from each other. However, when observed closely, it is clear that CM is positioned the furthest from GM and SM, while SM and GM are positioned relatively closer to one another. This seems to suggest that CM has a very different composition, while GM and SM appears to have comparable compositional properties (based on the NMR metabolites, see **Table 4.1**). The classification between the milk types is expected from the result of relative identification of the metabolites found in the NMR spectra (**Figure 4.1**), where CM is lacking more compounds compared to GM and SM as summarised on the Venn diagram (**Figure 4.2**). In terms of reproducibility, SM has the best reproducibility with the replicates positioned closely to each other, followed by GM and CM.

Following the score plot, the loading plot was generated in PCA to observe the correlation amongst NMR metabolites and their contribution to the PCs selection. In other words, the loading plots were used to determine the importance of the metabolites in classifying the three different milk types.

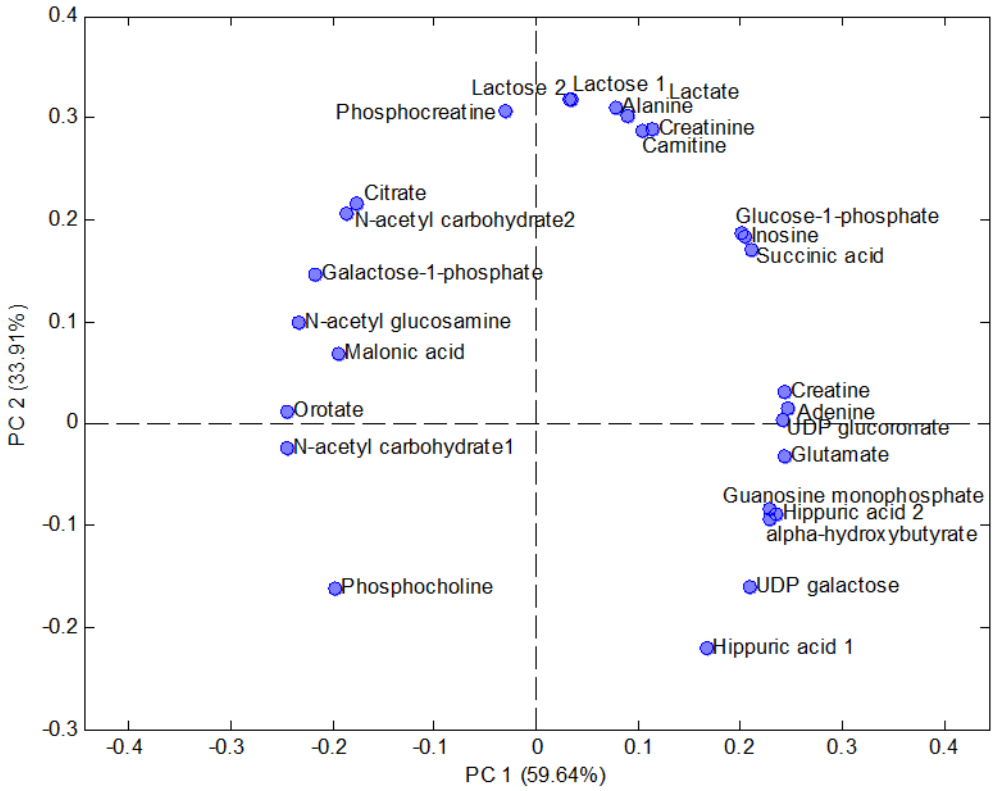

Figure 4.6. PCA loading plots for CM, GM, and SM metabolites

When the loadings variable is close to 1, it has strong positive correlation to the PC. On the other hand, if it is close to -1, the metabolites have strong negative correlation to the PC. Loadings close to 0 means the metabolites has a weak influence on the PC, specifically for the observed classification. To put it another way, the further the components are from the centre, the higher the discriminative power (Arcena, Kebede, Leong, Silcock, & Oey, 2020). As seen on **Figure 4.6**, most of the components are positioned away from the centre, indicating that a large number of these compounds has strong discriminative power. Even so, the metabolites positioned close to each other are found to have strong correlation.

When scores plot and loadings plot are overlaid on the same graph, a bi-plot is generated. Bi-plot shows the relationship between samples and metabolites. In the plot, the compounds positioned closer to a certain milk type is positively correlated toward that sample (or detected in higher amounts in that sample); while compounds that are farther away and projected in the opposite direction from a certain milk type are negatively correlation. To conclude**,** it is evident there are metabolites showing a strong relation to each types of milk and all the milks are different from each other (**Figure 4.7**).

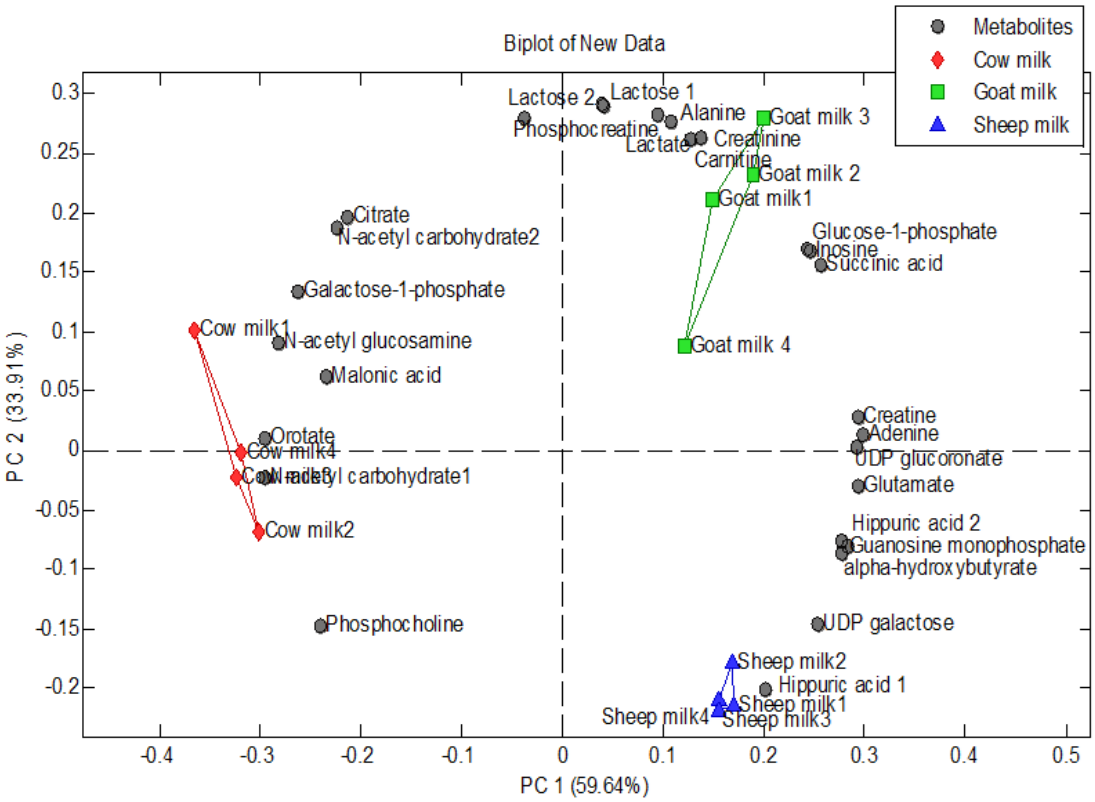

Figure 4.7. A PCA biplot illustrating the compound variance between NZ CM, GM and SM

In summary, the PCA has demonstrated its capability as a powerful unsupervised technique to explore the patterns and trends within the data. PCA showed a clear separation among the milk types without taking the sample information into account. As mentioned previously in literature, PCA is more powerful compared to supervised technique as it can simplify the complexity of high-dimensional data sets. Unlike supervised technique, it can find patterns without prior knowledge about the reference sample. PCA is used as initial screening tool, if the result from PCA is not good, there is no reason in performing supervised analysis into the data set.

In the current study, supervised chemometrics technique was applied into the data set to investigate the classification and identify discriminant compounds. In this case, PLS-DA was applied. As mentioned previously in literature (see **Section 2.7.2.2**), PLS-DA is a linear supervised clarification model that can predict and describe the class of the milk samples, whether it is CM, GM, or SM. It enables the selection of the most discriminative and predictive metabolites in the data that can aid the sample classification process. More information about the result of the PLS-DA is further discussed in **Section 4.4.3**.

4.4.3 Supervised PLS-DA of NZ Cow Milk (CM), Goat Milk (GM), and Sheep Milk (SM)

Following the result of PCA, further analysis was done with the use of PLS-DA. One of the main outputs of PLS-DA is a set of components called latent variables (LVs). LVs are the linear combination of the original variables, responsible to determine the optimal complexity of the model in PLS-DA (Ballabio & Consonni, 2013). A 'leave one out' cross validation was performed in this study to determine the optimal number of LVs needed to obtain the maximum amount of variance with the lowest noise or error.

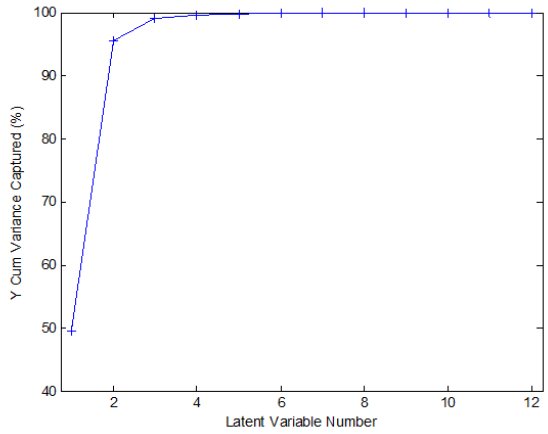

Figure 4.8. Cumulative Variance Graph – Characterization of NZ CM, GM and SM (PLS-DA)

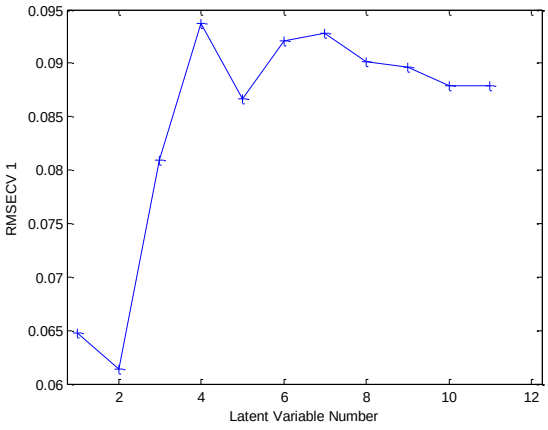

Figure 4.9. Root Mean Square Error of Cross Validation Graph - Characterization of NZ CM, GM, and SM (PLS-DA)

Based on cross validation (**Figure 4.8** and **Figure 4.9**), 2 LVs were chosen representing a higher cumulative variance while keeping the noise (root mean square error) to a minimum. The 2 LVs represent the cumulative variance of 97.33% and 96.04% of the X- and Y- variances, respectively. Once the number of LVs were chosen, a score plot (**Figure 4.10**) was generated.

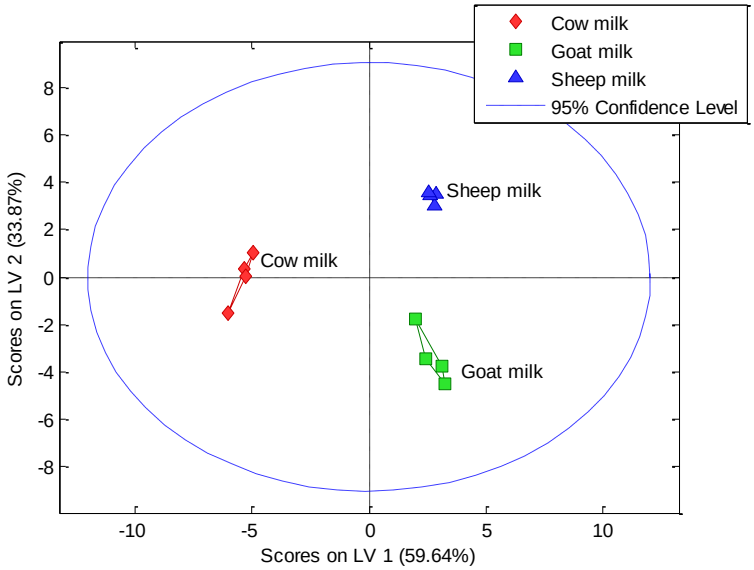

Figure 4.10. PLS-DA scores plot on the characterization of NZ CM, GM, and SM

In this study, the scores represent the coordinates of the three different milk samples in the L projection hyperspace. In another words, the score plot of PLS-DA is the result of the observation rows of X projected onto a hyperplane within the data, and it shows the covariance between X and Y-variables. Similar to the result from PCA (**Figure 4.5**), there are three clusters of milk samples (cow, goat, and sheep) with a clear separation between them, indicating that all the milk types are different from each other

To compliment the score plot, a loading plot was generated. The loadings must be non-zero for it to have contribution to the model (Ruiz-Perez & Narasimhan, 2018). In this case, the metabolites whose loadings are co-located from the centre of origin may be presumed to be correlated from one another.

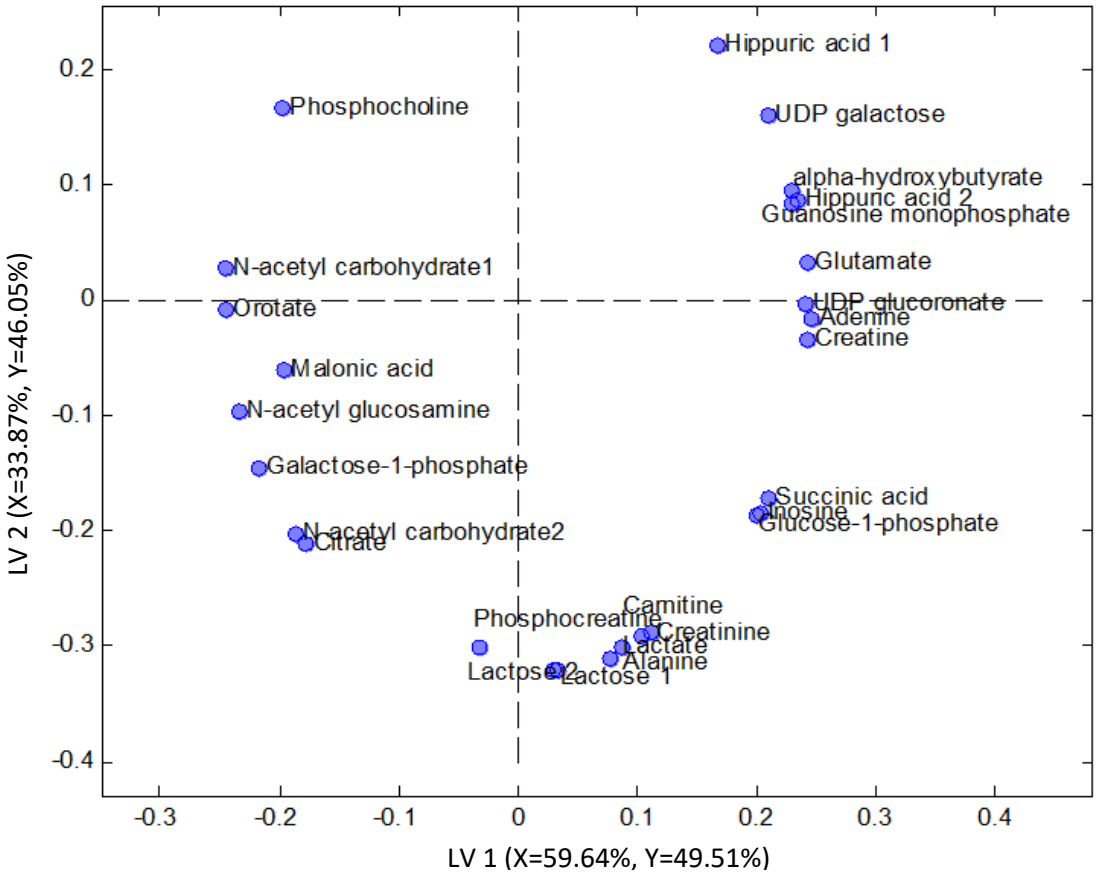

Figure 4.11. PLS-DA loading Plot for the characterization of NZ CM, GM, and SM

Due to the complementary nature of scores and loadings as clarification of the rows and columns of **X**, both can be used conjointly to generate a new biplot (**Figure 4.12**). In this study, the bi-plot is used to summarize how metabolites (X-variables) relate to each other as well as to the milk sample clusters (Y-variable). On the loading plot, the metabolites with loadings in a given position are presumed to contribute heavily towards the samples whose scores are found in a similar position in a scores plot (Worley & Powers, 2013). To put it another way,

the metabolites closest to the milk sample clusters are expected to have the highest contribution towards their class separation.

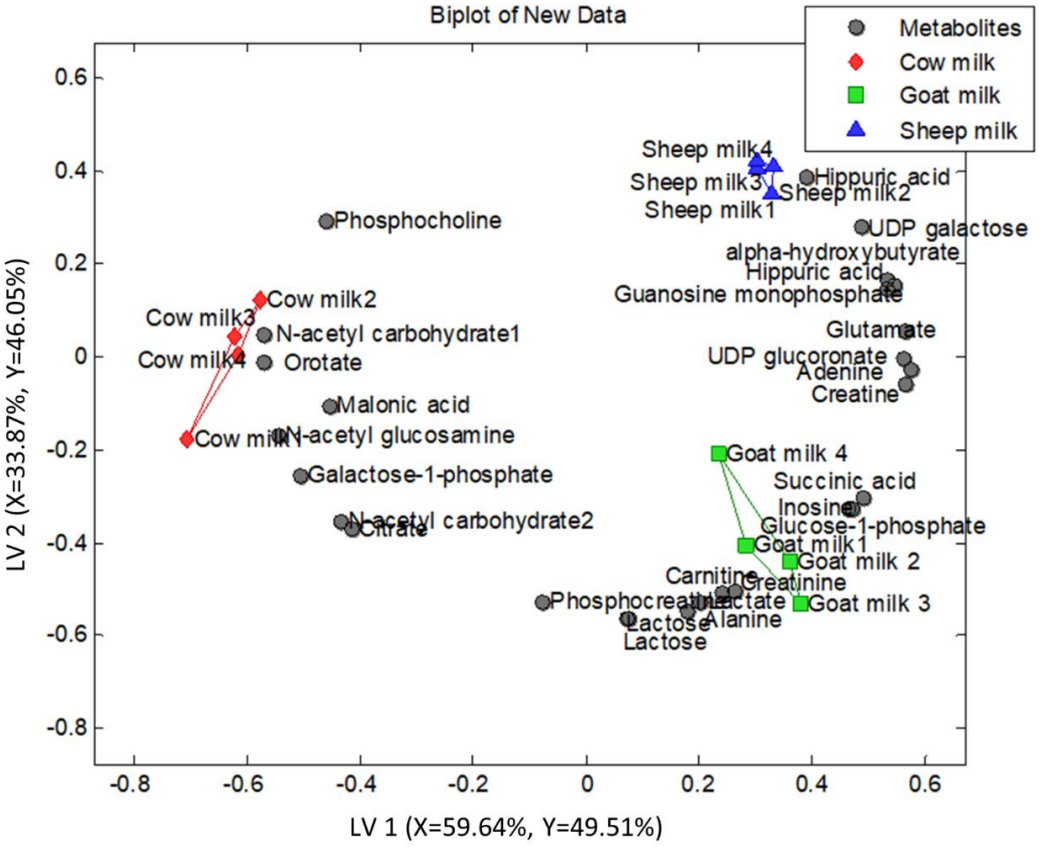

Figure 4.12. PLS-DA bi-plot on characterization of NZ CM, GM and SM

Overall, the use of PLS-DA in this study proved that it is a good method for milk characterization. PLS-DA was able to classify the milk types and identify the unique compounds that distinguished the milk from another. The model was capable of explaining the relationship between the metabolites and the milk types, whether they are positively correlated, negatively correlated, or uncorrelated. Even so, the bi-plot in PLS-DA was not straightforward in ranking the metabolites' importance towards the different milk types. It did not give the information about which metabolites can be used as discriminant markers to differentiate in between CM, GM and SM. As a result, further feature selection analysis was performed and followed by the calculation of VID coefficient.

4.4.4 Discriminant Markers Selection with VID and Interpretation

VID coefficients were calculated to identify which metabolites can be used as discriminant markers of one milk from another. Following the VID procedure, each compound was assigned with a coefficient between -1 and +1 in respect to the milk type. To determine the most

important compounds, only those with an absolute value of more than 0.700 were selected. All the discriminant compounds present in CM, GM, and SM are listed in decreasing order of VID coefficient, as shown in **Table 4.2**.

Table 4.2. Discriminant marker compounds selected for cow milk (CM)[1]

Cow Milk (CM)		
Identity	Chemical class	VID
Orotate	Carboxylic acid	0.98
N-acetyl carbohydrate1	Carbohydrates	0.98
N-acetyl glucosamine	Carbohydrates	0.94
Galactose-1-phosphate	Carbohydrates	0.87
Malonic acid	Dicarboxylic acid	0.79
Phosphocholine	Amino acid	0.78
N-acetyl carbohydrate2	Carbohydrates	0.75
Citrate	Tricarboxylic acid	0.71
Glucose-1-phosphate	Carbohydrates	-0.80
Inosine	Nucleosides	-0.81
UDP galactose	Carbohydrates	-0.84
Succinic acid	Dicarboxylic acid	-0.84
Alpha-hydroxybutyrate	Hydroxybutyric acid	-0.92
Guanosine monophosphate	Nucleotide	-0.92
Hippuric acid 2	Carboxylic acid	-0.94
UDP glucoronate	Carbohydrates	-0.97
Glutamate	Amino acid	-0.97
Creatine	Amino acid	-0.97
Adenine	Nucleobases	-0.99

[1]Only compounds with absolute VID coefficients of 0.700 or more are selected. Compounds are listed in decreasing order of VID coefficient. The positive VID coefficients indicate high presence of the compound while negative coefficient denote a low presence of the compound in milk.

Based on **Table 4.2**, CM had 8 positive discriminant compounds and 11 negative discriminant compounds. Compounds with a positive VID value includes, orotate, N-acetyl carbohydrate1, N-acetyl glucosamine, galactose-1-phosphate, malonic acid, phosphocholine, N-acetyl carbohydrate2, and citrate. If particular compounds have positive VID coefficient values, it indicates that those compounds are present in high amounts in CM samples in comparison to the other milk types. However, this also means that these compounds are present in lower amounts in other milk types. Compounds like glucose-1phosphate, inosine, succinic acid, UDP galactose, α-hydroxybutyrate, guanosine monophosphate, UDP glucoronate, creatine, glutamate, and adenine, which had negative VID values were also selected as discriminant markers. This indicate that these compounds are present in lower abundance in CM, when compared to GM and SM samples.

Following VID procedures, boxplots were used to depict the relative quantification of the selected discriminant markers in CM. Tukey's range test was performed to observe the significant differences in between the amount of discriminant markers present in the different milk types. Boxplots followed by the same letter indicates that there is no significant difference for in between the number of metabolites present in milk samples. On the other hand, boxplots followed by different letters indicate that there is a significant difference in between the number of metabolites present in the different milk types. The unit used for the relative quantification of each metabolites was p.d.u (procedure defined unit), in which it was based on the value of the internal standard (TSP) used in NMR spectrum.

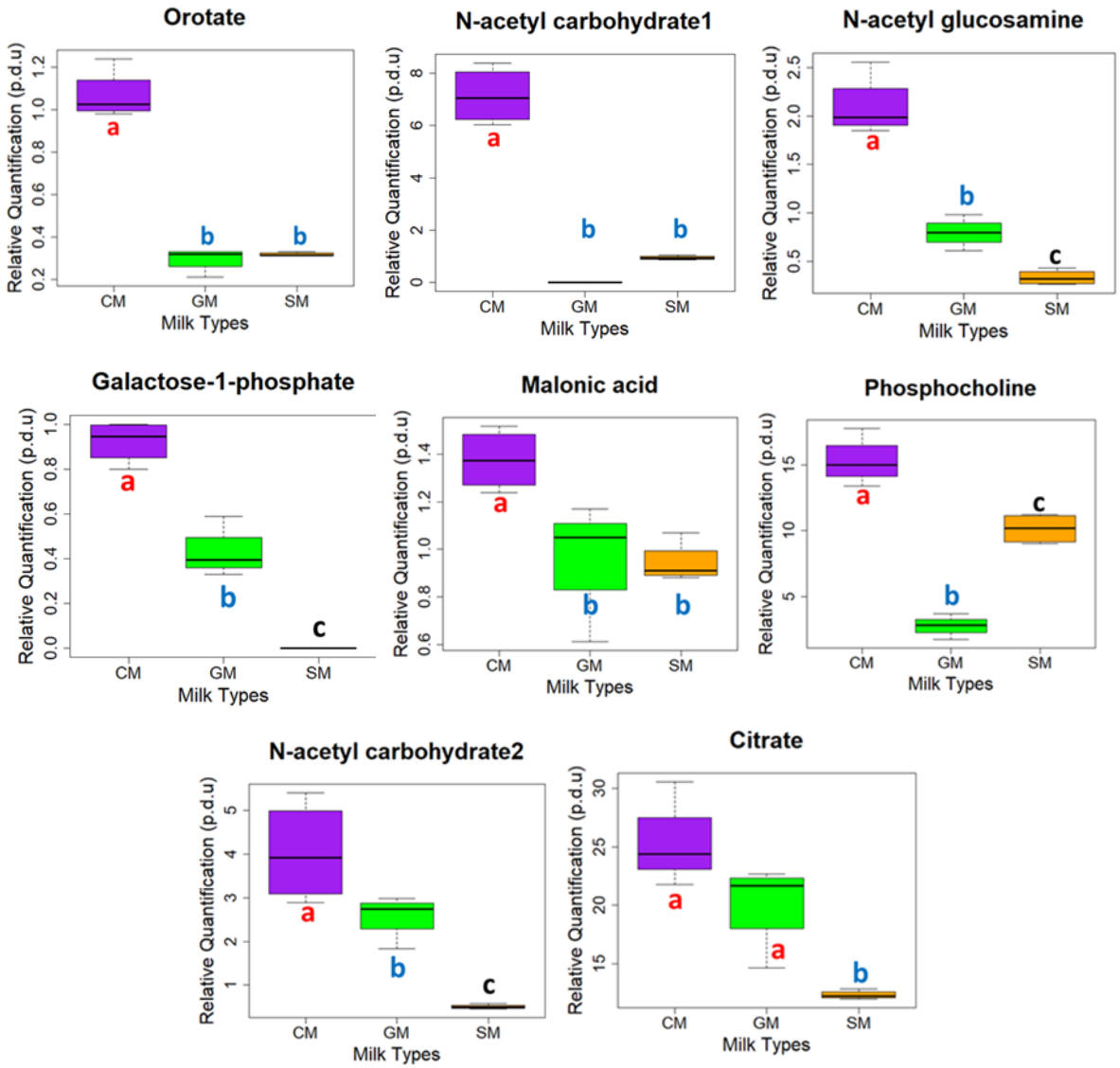

Figure 4.13. Positive discriminant compounds selected through VID procedures for cow milk (CM)

In the current study, orotate had the highest VID values as discriminant markers of CM. When analysed, the relative quantity of orotate was found to be significantly higher in CM compared to SM and GM (see **Figure 4.13**). Looking at past studies, similar results were found by Akalin and Gonc (1996), with orotic acid levels being highest in CM, followed by SM and GM. In another study, Wehrmueller, Jakob, and Ryffel (2008) also reported the same results, where CM was found to have the highest level of orotate followed by SM, and GM. Even though in the past studies by Akalin and Gonc (1996) and Wehrmueller et al. (2008) both measured the level of orotate in CM by HPLC instead of NMR, the trend of results was proven to be the same. Thus, NMR combined with advanced chemometrics is a good method in identification and quantification of milk metabolites constituent.

The compounds N-acetyl carbohydrates and N-acetyl glucosamine are, both oligosaccharides and are essential components of living organisms such as bacteria, plants, and animals (Mobli & Almond, 2007). In milk, they are considered as bioactive components that have prebiotic effects (Mijan, Lee, & Kwak, 2011). In the current study, N-acetyl carbohydrates and N-acetyl glucosamine were selected as discriminant compounds for CM. This was supported by the studies by Sundekilde, Larsen, et al. (2013) and Li, Yu, et al. (2017), where the presence of N-acetyl carbohydrates and N-acetyl glucosamine were reported in CM. However, they did not mention the relative quantification of the selected compounds present in CM in comparison to other milk types. In addition, N-acetyl carbohydrates was also selected as biomarkers for UHT and reconstituted milk (Cui et al., 2019).

Other discriminant markers (i.e., phosphocholine, malonic acid, and citrate) were also selected as their VID coefficients were greater than 0.700. Phosphocholine is the most abundant during the early stage of lactation and decreases as the lactation stage progresses (Artegoitia, Middleton, Harte, Campagna, & de Veth, 2014). This was supported by Klein et al. (2012) who reported that phosphocholine was selected as a biomarker to select healthy and stable cows for breeding purpose. For this reason, phosphocholine could be considered a good discriminant marker of CM. On the other hand, there are no reported literature about the relationship between malonic acid and CM. At the same time, although citrate was selected as a discriminant marker for CM, when Tukey's range test was performed on the relative quantification of citrate, the level of citrate in CM was not significant when compared to GM (see **Figure 4.13**). This was supported by the work of Peaker and Linzell (1975), where citrate levels in CM and GM were found to be similar.

Table 4.3. Discriminant marker compounds selected for goat milk (GM)[1]

Goat Milk		
Identity	Chemical class	VID
Creatinine	Amino acid	0.96
Carnitine	Fatty acid	0.95
Inosine	Nucleosides	0.95
Glucose-1-phosphate	Carbohydrates	0.94
Succinic acid	Dicarboxylic acid	0.93
Lactate	Carboxylic acid	0.93
Alanine	Amino acid	0.93
Lactose 2	Carbohydrates	0.84
Lactose 1	Carbohydrates	0.84
Phosphocholine	Amino acid	-0.89

[1]Only compounds with VID coefficients of 0.700 or higher are selected. Compounds are listed in decreasing order of VID coefficient. The positive VID coefficients indicate high presence of the compound while negative coefficient denote a low presence of the compound in the sample.

Based on **Table 4.3**, GM had 9 positive discriminant and 1 negative discriminant compounds. Compounds such as creatinine, carnitine, inosine, glucose-1-phosphate, succinic acid, lactate, alanine, and lactose were present in high amount, shown by the positive VID values, while phosphocholine was present in a low amount in GM samples, shown by their negative VID values.

Next, different boxplots (**Figure 4.14**) were constructed to visualise the relative quantification of the selected discriminant markers in GM. To see the significant differences in between the amount of discriminant markers present in the milk type, Tukey's range test was performed. Boxplots followed by the same letter indicates that there is no significant difference for in between the number of metabolites present in milk samples. On the other hand, boxplots followed by different letters indicate that there is a significant difference in between the number of metabolites present in the different milk types.

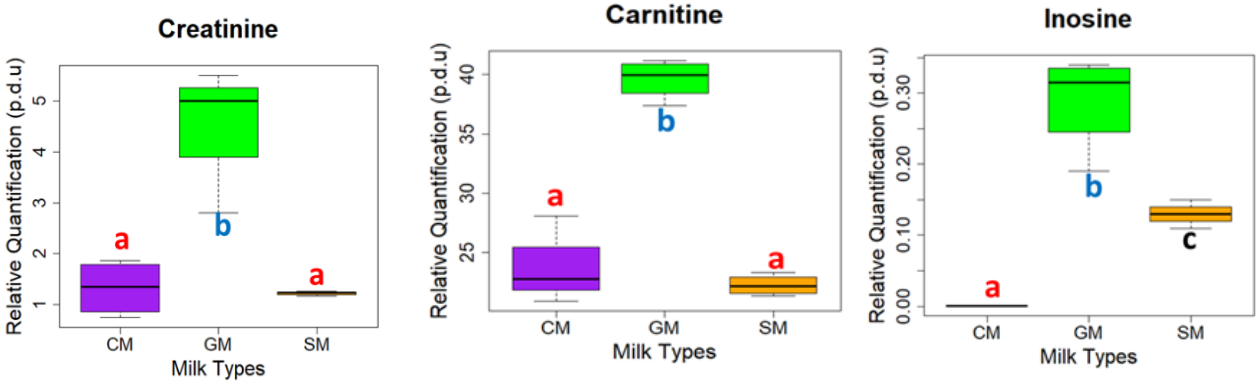

Figure 4.14. Positive discriminant compounds selected through VID procedures for goat milk (GM)

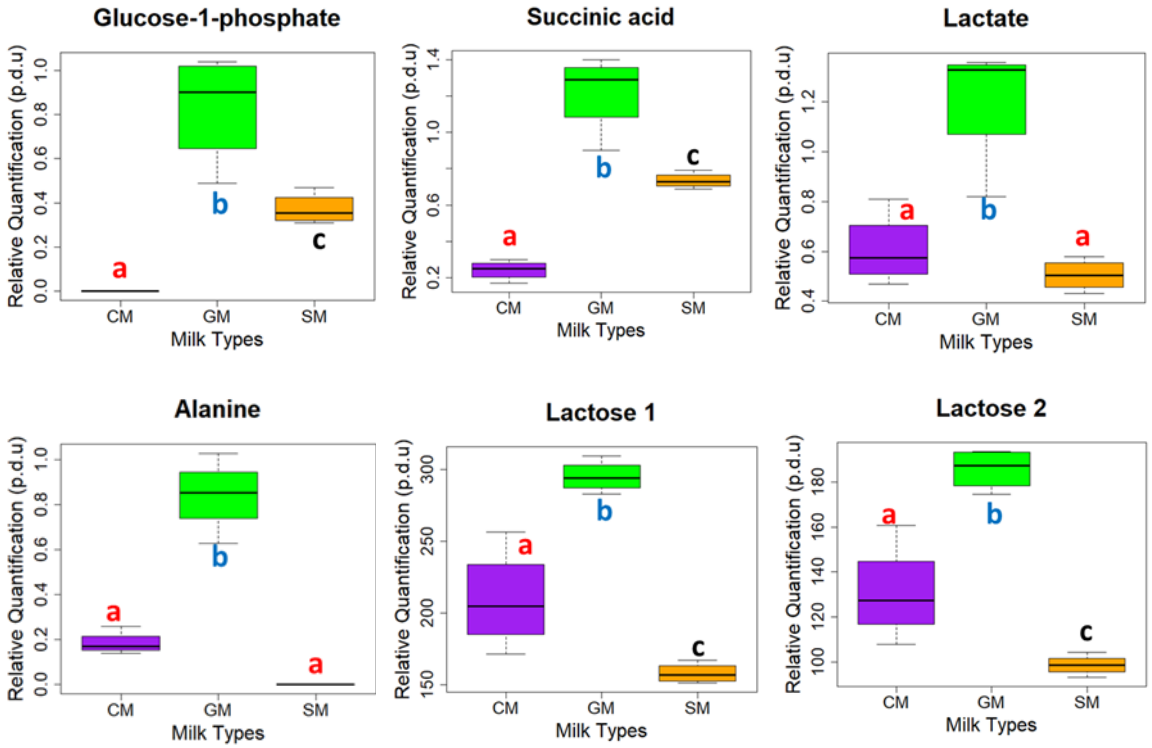

Figure 4.14. Positive discriminant compounds selected through VID procedures for goat milk (GM) (continued)

For GM, creatinine had the highest VID value and served as discriminant markers of GM. Creatinine was found to be the highest in GM, while there was no significant difference in CM and SM. Creatinine is a by-product secreted by goat for their normal muscle metabolism. Creatinine is also part of the metabolites used to estimate the concentration of urine in several ruminants (Dos Santos et al., 2018). Similar to creatinine, carnitine was also found to be the highest in GM, while there was no significant difference in the levels of carnitine in either CM or SM. Carnitine is part of the NPN fraction in milk commonly used in the development of neonate (Hoppel, 2003). In ruminants, carnitine is said to have an important role in facilitating the transport of fatty acids into mitochondrial matrix for oxidation during thermogenesis and ketogenesis process (Y. W. Park et al., 2007). According to published literature, the amount of carnitine present in ruminant milk is always lower than the amount in human milk (Penn, Dolderer, & Schmidt-Sommerfeld, 1987). Overall, the published literature on creatinine and carnitine contents of CM, GM, and SM are scarce. Nevertheless, based on the fingerprinting study of NZ GM by Sanchez et al. (2020), both creatinine and carnitine are present in NZ GM.

Inosine levels in GM was found to be significantly different from those in SM, while no inosine was detected in CM. This contradicts the result of prior studies about the presence and levels of nucleosides and nucleotides in SM and GM. According to Plakantara, Michaelidou,

Polychroniadou, Menexes, and Alichanidis (2010), the level of inosine of SM and GM was comparable, and no significant differences were found between the two. The findings in this study could be because the presence of inosine in milk is affected by the lactation stage of the species. Inosines are usually present at highest concentration during the early lactation stage of the mammals, but this amount decreases as the lactation days goes on (Holmes-McNary, Cheng, Mar, Fussell, & Zeisel, 1996). However, lactation stage could not be considered as a factor in the current study since it is unclear what lactation stages were in existence when the two milk types (GM and SM) were obtained.

Moreover, the fact that there was no inosine found in CM in the current study is supported by the findings of past ^{1}H-NMR studies. Sundekilde et al. (2014) reported that there is no inosine in the milk metabolome of Swedish dairy cows. Additionally, inosine also was not reported in the metabolites found in the study of CM authenticity (Q. Li, Yu, et al., 2017).

There is no information about the relationship between glucose-1-phosphate and GM in the published literature. Nonetheless, the amount of glucose-1-phosphate in GM was the highest, and the amount is significantly different in SM and CM. Significant differences exist in the levels of succinic acid. Referring to past studies, there has not been investigation about the presence of succinic acid in CM, GM, and SM. Nevertheless, succinic acid is known to be present in most ruminant milks (Caboni et al., 2016; Dursun, Güler, & Şekerli, 2017). Similar to glucose-1-phosphate, there was no information in published literature on the level of lactate present in GM. Even so, lactate is generally present in ruminant milk as their presence is caused by the fermentation of lactose (Garrote, Abraham, & Rumbo, 2015).

Alanine was found to be the highest in GM, followed by CM, while there was no alanine found in SM. The fact that alanine was not found in SM contradicts the findings from the past studies. Claeys et al. (2014) found that the levels of alanine in SM is twice the amounts present in GM, and almost three times the amount in CM. However, in another study, alanine was found to be the highest in GM, while the level of alanine in CM and SM was reported to be similar (Rafiq et al., 2016). Overall, the different findings about the presence of alanine in milk could have been due to factors such as genetics, breed, diet, and other seasonal variation that influenced the free amino acid contents in milk (Alhussien & Dang, 2018). Other than that, the difference could also be caused by the utilisation of different analytical methods used to measure milk composition (i.e. the result from UPLC will be different from the result from NMR even if the same sample are analysed by both techniques) (Zhu, Kebede, Chen, McComb, & Frew, 2020c).

Lastly, lactose was found to be highest in GM followed by CM and SM (see ***Figure 4.14***). In fact, here was a significant difference between the amounts of lactose in the three milk types. Yet, the outcome of the present study does not match with past studies on the lactose content of CM, GM, and SM. According to the study on the physicochemical characteristic of GM and SM, lactose is highest in SM (4.9%), followed by CM (4.7%) and GM (4.1%) (Park et al., 2007). In a study by Claeys et al. (2014), SM had an average lactose level of 5.9%, while CM and SM had an average of 5.6% and 5.0% respectively. More recently, C.F. Balthazar et al. (2017) also found SM to have the highest level of lactose (with 4.8%), while CM had 4.7%, and GM had 4.1%. When comparing the literature with the result of the current study, these implies that there might have been factors that influenced the lactose content of the tested milk. One of the possible factors is lactation stage. At the end of the lactation stage of mammals, the volume of milk is reduced, and this results in changes in milk composition (Panthi, Jordan, Kelly, & Sheehan, 2017). Lactose content in ruminants typically decrease during the late stage of lactation. Accordingly, the milk samples used in this study might have been collected at different lactation stages of the animals. Even so, as mentioned previously, Lactation stage could not be considered as a factor in this study, and it is unclear what lactation stages were in existence when the two milk types of SM and GM were obtained. Also, different analytical technique may have been used during the measurement of lactose content in the different milk types, and this too may have led to the differences in the results.

Table 4.4. Discriminant marker compounds selected for sheep milk (SM)[1]

Sheep Milk		
Identity	Chemical class	VID
Hippuric acid 1	Carboxylic acid	0.94
UDP galactose	Carbohydrates	0.90
Alpha-hydroxybutyrate	Hydroxybutyric acid	0.80
Hippuric acid 2	Carboxylic acid	0.79
Guanosine monophosphate	Nucleotide	0.77
Phosphocreatine	Amino acid	-0.80
N-acetyl glucosamine	Carbohydrates	-0.81
Galactose-1-phosphate	Carbohydrates	-0.89
Citrate	Tricarboxylic acid	-0.94
N-acetyl carbohydrate2	Carbohydrates	-0.95

[1]Only compounds with VID coefficients of 0.700 or higher are selected. Compounds are listed in decreasing order of VID coefficient. The positive VID coefficients indicate high presence of the compound while negative coefficient denote a low presence of the compound in the sample.

Based on **Table 4.4**, SM has 5 positive and 5 negative discriminant compounds. The compounds with positive VID values were hippuric acid, UDP galactose, α-hydroxybutyrate,

guanosine monophosphate, and glutamate. The compounds with negative VID values included phosphocreatine, N-acetyl glucosamine, galactose-1-phosphate, citrate, and N-acetyl carbohydrate 2.

Boxplots were constructed to visualise the relative quantification of the selected discriminant markers in SM (**Figure 4.15**). Boxplots followed by the same letter indicates that there is no significant difference in values recorded for the metabolites present in milk samples. On the other hand, boxplots followed by different letters indicate that there is a significant difference in between the number of metabolites present.

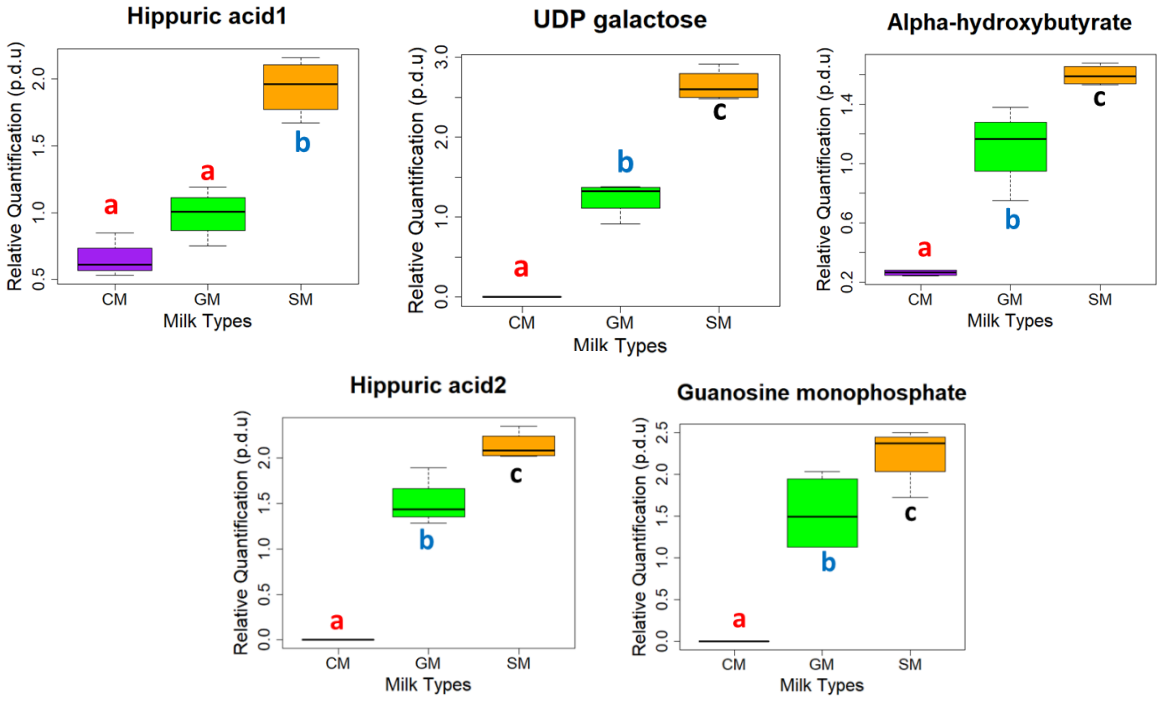

Figure 4.15. Positive discriminant compounds selected through VID procedures for sheep milk (SM)

Based on **Table 4.4**, hippuric acid1 had the highest VID coefficient and chosen as one of the discriminant compounds for SM. In general, hippuric acid is an organic compound that served as one of the constituents of NPN fractions in milk. According to Carpio et al. (2013), hippuric acid can be used as a biomarkers for CM and GM from different feeding regimen. In the current study, the hippuric acid content was found to be the highest in SM, followed by GM, and CM when analysed by NMR. This was supported by findings reported in earlier studies. Hornickova, Dragounová, Hejtmánková, Michlova, and Hejtmánková (2015) reported higher hippuric acid content in SM than GM, while there was no correlation between the time of sample collection at different lactation stages with the hippuric acid levels in SM and GM analysed by HPLC. In another study, the levels of hippuric acid in GM was found to be higher

than CM when analysed by CE (Carpio, Rodríguez-Estévez, Sánchez-Rodríguez, Arce, & Valcárcel, 2010). To sum up, there is no correlation between the levels of hippuric acid in different types of milk and the different analytical technique used to measure its levels.

As a compound that has the second highest VID compounds in SM, UDP galactose was the highest in SM, followed by GM, while this compound was not detected in CM (see **Figure 4.15**). In milk, UDP galactose is formed by the reaction of galactose-1-phosphate and UDP-glucose (Mohammad, Hadsell, & Haymond, 2012). In the literature, the information regarding the level of UDP-galactose in milk is very rare and there has only been one publication about UDP-galactose content in ruminant milks. According to Gil and Sanchez-Medina (1981), UDP-galactose is present in CM, GM, and SM and that it decreases as the lactation stage progresses.

α-hydroxybutyrate is typically used as a biomarker of lipid oxidation and insulin resistance (Sarı, Esen, Yıldırım, Pilten, & Aydın, 2019). Nonetheless, in this study α-hydroxybutyrate was found to be one of the discriminant markers for SM. No previous study had investigated the presence of α-hydroxybutyrate in milk products, instead they investigated the presence of β-hydroxybutyrate.

Guanosine monophosphate (GMP) also known as guanylic acid is one of the purines found in milk products. GMP was selected as one of the discriminant compounds for SM due to its high VID coefficient (see **Figure 4.15**). In the literature, Gil and Sanchez-Medina (1981), using an enzymatic analysis method had previously reported the presence of GMP in SM but not in CM and GM. Contrary to that findings, GMP was found in CM using reversed-phase HPLC (Tiemeyer, Stohrer, & Giesecke, 1984). In another study, GMP was found in both infant formula and CM using ^{1}H-NMR metabolomics method (Zhao et al., 2017). GMP was also found as one of the available metabolites in NZ GM using ^{1}H-NMR metabolomics method (Sanchez et al., 2020). Although, GMP was known to be present in different milk types, there was no information regarding the level of GMP available in each milk from the literature. Furthermore, different analytical technique used to detect the presence of GMP found in the literature yielded different result. This indicates that the application of certain analytical techniques has a crucial role in detecting the metabolites.

4.4.5 ConclusionChapter 4 and Next Steps

Based on the result of the NMR spectra (**Figure 4.1**), different metabolites are present in different milk types. CM was found to have 17 metabolites, SM with 23 metabolites, and GM

with 24 metabolites (**Figure 4.2**). Following this discovery, advanced chemometrics techniques (unsupervised PCA and supervised PLS-DA) were used to explore the data to classify the milk samples and explain the correlation between the metabolites and the milk samples.

Next, the VID coefficient was selected to determine the discriminant compounds for three different milk types. Out of the three, CM had the highest number of discriminant compounds followed by GM, and SM (see **Table 4.2**, **Table 4.3**, and **Table 4.4**). Accordingly, boxplots were constructed for the selected discriminant compounds followed by Tukey's range test in order to know whether there are significant differences in the discriminant compounds among milk types (CM, GM, and SM).

In order to achieve the objective of the study, which is to detect adulteration of NZ CM, GM, and SM, Chapter 4 was used as the first step to identify metabolites characterising each different milk types. In the next step, it is important to identify and characterise metabolites that can be used as markers of adulterations. This therefore is the focus of Chapter 5.

Chapter 5 . Detection of Adulteration of New Zealand Goat and Sheep Milk with Cow's Milk using NMR Spectroscopy and Chemometrics

5.1 Introduction

Different analytical techniques have been utilised for detection of milk adulteration and milk authenticity analysis. Scano et al. (2014) had used GC-MS based metabolomic approach to characterise the metabolite composition of CM and GM. Angelopoulou et al. (2015) had developed an optical biosensor to rapidly detect CM in GM. In another study, NIR Raman spectroscopy was used in detecting the presence of urea in milk (Khan et al., 2015). More recently, NMR-based metabolomics was used as tools for metabolite profiling in CM, GM, and soy milk (Li, Yu, et al., 2017). The combination of 1D and 2D NMR analysis was also used for identification of biomarkers for reconstituted milk (Cui et al., 2019).

Based on these studies, NMR-based metabolomics approach was proven to be powerful in metabolites and components analysis. Compared with most other approaches, NMR has a simpler analytical procedure, as well as less organic solvent consumption. These make NMR as an ideal technique in detecting adulterations and milk authentication. Even so, there has not been much study regarding the use of NMR on the detection of milk adulteration. Additionally, NZ GM and SM are considered as high-value products. The price of GM and SM is 2–3x the price of CM in NZ. These facts put NZ dairy products at risk of fraudulent activities such as milk adulteration or counterfeiting. Therefore, it is interesting to explore the potential of NMR-based metabolomics in the detection of adulteration in NZ milks.

In **Chapter 4**, the metabolite composition of NZ CM, GM, and SM powders were successfully characterized using an untargeted NMR-based metabolomics combined with an advanced chemometrics approach. Following advanced chemometrics, the VID coefficients of each metabolites were calculated to identify metabolites that can discriminate the milk samples.

In this chapter, the potential of NMR fingerprinting technique will be employed to detect the adulteration in NZ GM and SM samples that have been adulterated with different concentrations of CM (1%, 2%, 4%, and 8%). Advanced chemometrics and feature selection (e.g. VID) were used to identify markers of adulterations in these high value dairy products.

5.2 Materials and Method

The details regarding the materials and method used in the study are as follows.

5.2.1 Samples and Reagents

The samples used in this study are GM and SM that have been adulterated with 1%, 2%, 4%, and 8% of CM. A list of the reagents used in the study was previously mentioned in **Section 4.2.2.**

^{1}H-NMR analysis

5.2.2 Sample Extraction and Preparations

Detailed sample extraction and preparation steps was mentioned earlier in **Section 4.2.3.**

5.2.3 ^{1}H-NMR Experiments

The settings for the NMR apparatus was mentioned previously in **Section 4.2.4.**

5.3 Multivariate Data Analysis

In the present study, both supervised (PLS-R) and unsupervised (PCA) multivariate data analysis (MVDA) were employed to obtain a model to detect adulteration of NZ GM and SM milk with different concentrations of CM (1%, 2%, 4% and 8%). PLS-R and PCA were used as they are the most appropriate technique for a large data sets generated by NMR. They have the capability to effectively summarise the relationship between the adulterated samples and identify the metabolites characterising a certain group or trends in the milk samples.

5.3.1 Unsupervised Principal Component Analysis (PCA)

Unsupervised PCA was performed in the study to compare the adulterated milk samples with each other and to visualise trends and clusters based on the level of fortification. More details regarding the software were mentioned in **Section 4.3.1.**

5.3.2 Supervised Partial Least Square Regression Analysis (PLS-R)

Supervised PLS-R was used to investigate the trend of the metabolite profile as a function of different adulteration concentrations, to know which metabolites were increasing or decreasing as the percentages of adulterants increases. For PLS-R, the datasets were loaded into the SOLO software in which, the metabolites were considered as the X-variable and the concentration of adulterants (1%, 2%, 4%, and 8% CM) were considered as the continuous Y-variable. Same with PCA and PLS-DA, a 'leave one out' cross validation was applied, and cumulative variance graph and the root mean square error of cross validation graph were constructed. To avoid the risks of overfitting, the optimum number of latent variables (LVs) were selected based on cross-

validation. The criterion for the selected LVs is that it must explain the maximum variance of the data set with the lowest possible error (RMSECV) (Kebede, Ting, Eyres, & Oey, 2020).

5.3.3 Markers Selection with VID

To identify metabolites that increase or decrease as a function of the adulteration concentration, variable identification (VID) coefficients were calculated. In this current study, only compounds with VID value higher than 0.700 were selected. For each discriminant markers, a box and whisker plot were constructed individually using RStudio Desktop (Version 1.3.1093, RStudio Team, Boston MA, USA). Each box plot displayed the five-number summary of the relative amounts of compounds present in the unadulterated milk and the milk with different concentration of adulterants (1%, 2%, 4%, and 8% CM). The five-number summary includes the minimum, first quartile, median, third quartile, and the maximum relative amount of the selected discriminant compounds. For further analysis, SPSS Software Version 24 (IBM Corporation, New York, New York, United States) was employed to perform Tukey's range test for significant different testing ($p < 0.05$) in NZ CM, GM, and SM samples (to confirm the discriminative potential of the metabolites selected with VID).

5.4 Results and Discussion

The results and discussion sections for the detection of adulteration of GM and SM with different concentration of CM are divided into two subsections as follows:

5.4.1 Adulteration of New Zealand Goat Milk (GM) with Different Concentration of Cow Milk (CM)

The result for the detection of adulteration of NZ GM with CM is explained in the following sections:

5.4.1.1 Result from 1H-NMR Spectra of Adulterated GM

A representative NMR spectrum of NZ GM adulterated with different concentration of CM is shown in **Figure 5.1**. The spectra were carefully inspected to ensure the information regarding the chemical shifts is accurate given the complexity of the identification process.

Similar to the discussion in **Section 4.4.1**, the spectra of the adulterated GM can be divided into three spectral regions. First is the aliphatic region covering amino acids, carboxylic acid, and carbohydrates ranges from 0 to 3.5 ppm. Second is the sugar region consisting of simple and complex carbohydrates ranges from 3.5 to 6.0 ppm. The last region is dedicated to the aromatic class including carboxylic acids, nucleobases, and nucleotides ranges from 6.5 to 8.5 ppm.

The NMR spectra shown on **Figure 5.1** are also comparable to the study findings by Sanchez et al. (2020), a study focused on the optimization of NMR and GC-MS technique in the characterization of GM powder. Similar to the current study, 3 different spectral regions including aliphatic, sugars, and aromatic regions were identified. High-intensity signals such as citrate and lactose were found which matches the result of this study. The current findings is supported by the previous study by Li, Yu, et al. (2017), where they found the linear relationship between the adulteration proportion of GM with CM and the content of characteristic variables including lactose, citrate, creatinine, and carnitine.

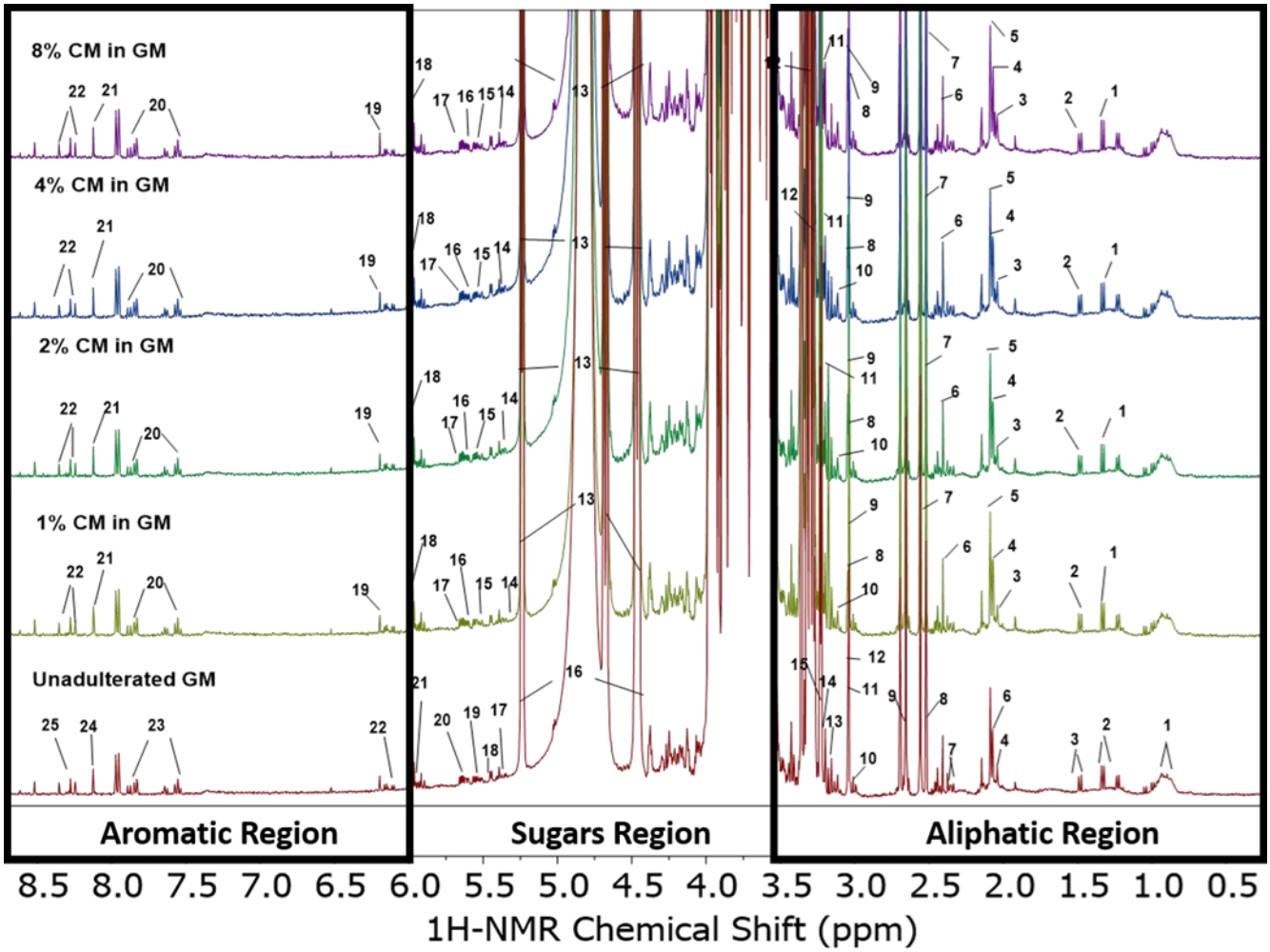

Figure 5.1. NMR Spectra of NZ GM adulterated with different concentration of CM (1%, 2%, 4%, and 8%)

In this chapter, the data analysis was performed only on discriminant compounds selected for the different milk types in chapter 4. This includes a total 22 metabolites, comprising 7 carbohydrates, 6 amino acids, 6 carboxylic acid, 2 nucleosides, 1 nucleotide, and 1 nucleobase. Details regarding its identity, chemical shifts, multiplicity, assignment, and the type of compounds found in adulterated GM can be seen in **Table 5.1**. On the other hand, the details regarding the compounds found in unadulterated GM can be seen in **Table 4.1**. To study the metabolite changes of GM as a result of adulteration, advanced chemometrics were performed. Further analysis was also employed to determine potential markers, followed by the explanation regarding the importance of the selected milk metabolites as markers of adulteration in GM.

Table 5.1. Metabolites assignment from [1]H-NMR spectra of NZ GM adulterated with different concentration of CM (1%, 2%, 4%, and 8%)

Peak	Identity	δ_{1H} (multiplicity)[1]	Assignment	Type of Compounds	Presence	Ref[2]
1	Lactate	1.33(d)	CH_3	Carboxylic acid	All Milk	[1],[5]
2	Alanine	1.48(d)	CH_3	Amino Acid	All Milk	[2],[4]
3	N-acetyl glucosamine	2.04(s)	CH_3	Carbohydrate	All Milk	[6],[7]
4	N-acetyl carbohydrate2	2.07(s)	CH_3	Carbohydrate	All Milk	[3]
5	Glutamate	2.09(d)	β-CH_2	Amino Acid	All Milk	[3]
6	Succinic acid	2.41(s)	CH_2	Carboxylic acid	All Milk	[1]
7	Citrate	2.54(d)	CH_2	Tricarboxylic acid	All Milk	[4]
8	Phosphocreatine	3.04(s)	CH_3	Amino Acid	All Milk	[1]
9	Creatinine	3.05(s)	CH_3	Amino Acid	All Milk	[5]
10	Malonic acid	3.12(s)	CH_2	Dicarboxylic acid	1%,2%,4% CM/GM	[2],[3]
11	Phosphocholine	3.20(s)	CH_3	Amino Acid	All Milk	[1]
12	Carnitine	3.23(d)	CH_3	Amino Acid	4%,8% CM/GM	[7]
13	Lactose	4.68(d)	CH	Carbohydrate	All Milk	[1],[3]
		5.24(d)	CH	Carbohydrate	All Milk	[1],[4]
14	Galactose-1-phosphate	5.36(dd)	CH	Carbohydrate	All Milk	[1]
15	Glucose-1-phosphate	5.55(dd)	CH	Carbohydrate	All Milk	[6],[7]
16	UDP glucoronate	5.61(dd)	CH	Carbohydrate	All Milk	[2]
17	UDP galactose	5.65(dd)	CH	Carbohydrate	All Milk	[1],[7]
18	Guanosine monophosphate	5.97(s)	CH	Nucleotides	All Milk	[1],[7]
19	Orotate	6.2(s)	CH	Carboxylic acid	All Milk	[2],[3],[6]
20	Hippuric acid	7.56(m)	CH_2	Carboxylic acid	All Milk	[3],[5]
		7.85(m)	CH_2	Carboxylic acid	All Milk	[1]
21	Adenine	8.12(s)	CH	Nucleobases	All Milk	[2],[3]
22	Inosine	8.25(s)	CH	Nucleosides	All Milk	[1]
		8.35(s)	CH	Nucleosides	All Milk	[1]

[1]Multiplicity: s=singlet; d=doublet; t=triplet; dd=double of doublets; m=multiplets.

[2]References: [1]=Human Metabolome Database (HMDB; http://hmdb.ca/); [2]=(Klein et al., 2010); [3]=(Sundekilde, Larsen, et al., 2013), [4]= (Sundekilde, Poulsen, Larsen, & Bertram, 2013), [5]=(Li, Yu, et al., 2017),[6]=(Zhao et al., 2017),[7]=(Sanchez et al., 2020)

5.4.1.2 Detecting the Adulteration using Chemometrics with Principal Component Analysis (PCA)

Unsupervised PCA analysis was performed to explore the trend and patterns due to the systematic adulteration of GM with different amount of CM and to check for outliers. Although the PCA results weren't shown, no outliers were detected in PCA. PCA indicated a clear change in the metabolite profile of GM samples as a function of adulteration. Consecutively, a supervised chemometrics technique was applied to further investigate the metabolite changes in milk samples and identify discriminant compounds that can be a candidate marker of adulteration. Hence, PLS-R is applied in the next section.

5.4.1.3 Detecting the Adulteration using Chemometrics with Partial Least Square Regression Analysis (PLS-R)

Different from PCA, PLS-R is a method for relating two data matrices, X and Y not only by a linear multivariate model but also by regression. PLS-R has the capability to analyse data with many collinear, noisy, and even incomplete data sets in both X and Y (Wold et al., 2001).

In the present study, PLS-R method was used to study the continuous changes in the adulteration of GM with different concentration of CM (1%, 2%, 4%, and 8%). The X variable represents the NMR metabolites and Y variable represents the concentration of the adulterated milk samples. A cross validation was performed to avoid overfitting of the data set and select an optimum number of latent variable (LVs), where a maximum variance is reached while keeping the noise at a minimum. These can be seen on **Figure 5.2** and **Figure 5.3** below.

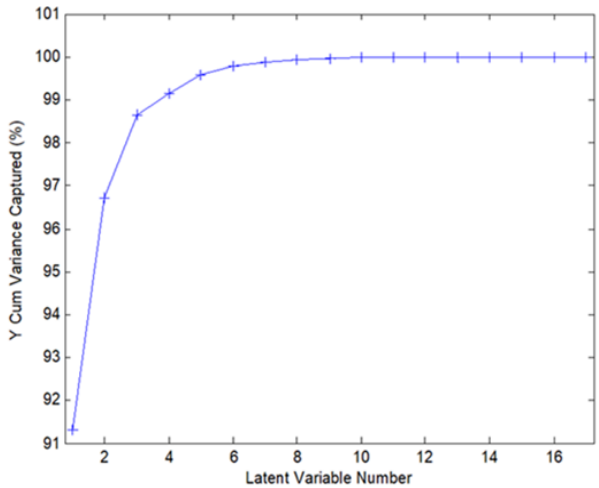

Figure 5.2. Cumulative Variance Graph - GM adulterated with CM (1%, 2%, 4%, and 8%) (PLSR)

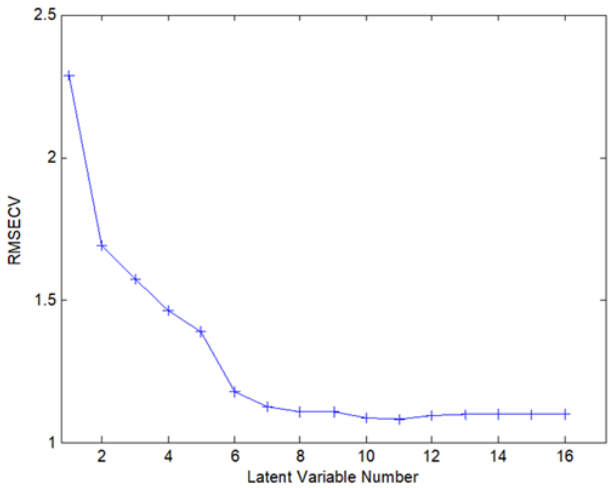

Figure 5.3. Root Mean Square Error of Cross Validation Graph - GM adulterated with CM (1%, 2%, 4%, and 8%) (PLS-R)

Based on the cross validation, 6 LVs were chosen. The 6 LVs represent 69.68% and 99.79% of the cumulative X and Y variance, respectively. The first 2 LVs represent the highest cumulative variance. Thus, the score plot, loadings plot, and bi-plot were constructed on the basis of LV1 and LV2. The score plot and loadings plot for adulterated GM can be seen on appendix (**Figure A.1** and **Figure A.2**), while the bi-plot (**Figure 5.4**) can be seen below.

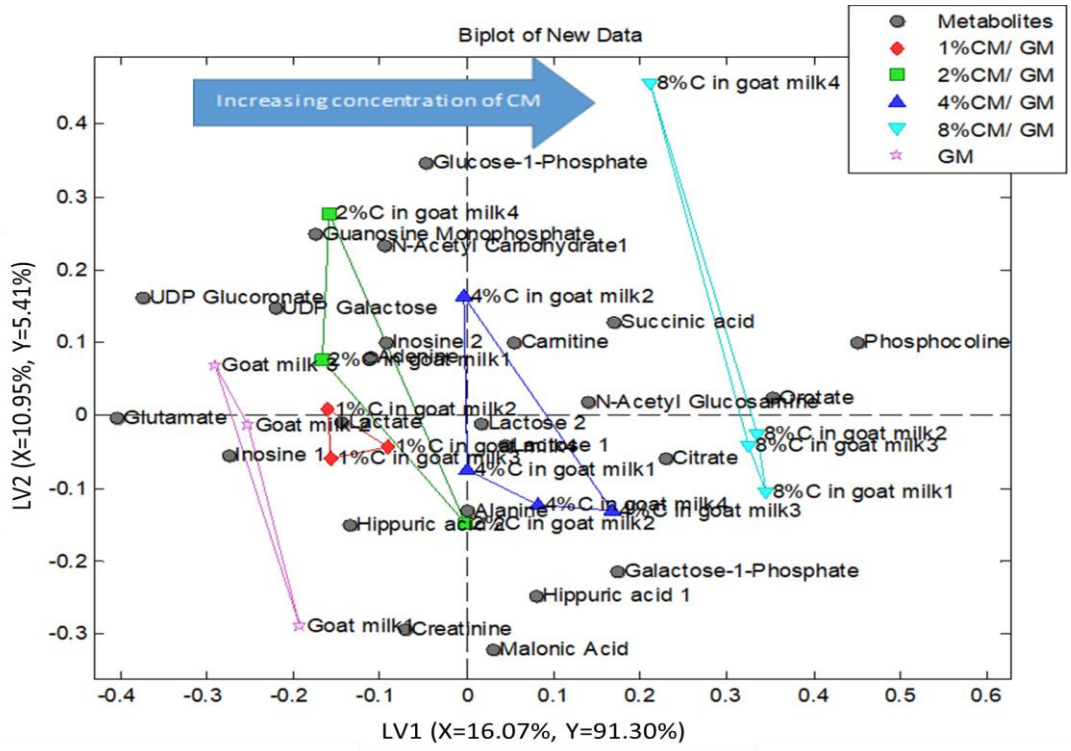

Figure 5.4. PLS-R bi-plot showing the change in the NMR metabolite profile of GM due to the adulteration with different concentration of CM (1%, 2%, 4%, and 8%)

Based on the bi-plot (**Figure 5.4**), there is a clear change in the metabolite profile of the GM adulterated with different concentration of CM (1%, 2%, 4%, and 8%). There is a trend from left to the right side of the bi-plot along with the increase of adulteration concentrations. The unadulterated GM is positioned on the far left, followed by GM adulterated with 1% of CM, GM adulterated with 2% of CM, GM adulterated with 4% of CM, and lastly GM adulterated with 8% of CM. The further the milk is positioned from the unadulterated sample, the more different its metabolic profile is.

The relationship between the metabolites and the level of adulteration can be interpreted based on the position of the metabolites on the bi-plot. Loadings located near each other are typically positively correlated and if they are located opposite each other they are negatively correlated.

To conclude, the developed PLS-R model enabled a successful classification and separation of different adulteration levels, even at a very small concentration of CM. Using the bi-plots, it was possible to visually describe the relationship between the metabolites and the level of adulteration. However, it was not possible to rank the metabolites based on their importance in explaining the classification and trend observed on the bi-plot. For this reason, further analysis to identify the discriminant markers of adulterated and unadulterated GM was performed by calculating VID coefficients.

5.4.1.4 Discriminant Markers Selection of Adulterated Goat Milk (GM)

The VID coefficients were calculated to identify metabolites that can be used as candidate markers of GM adulterations with CM. Following the VID procedure, each compound was assigned with a coefficient between -1 and +1 in each sample class. To determine the most important compounds, only those with an absolute value of more than 0.700 were selected (**Table 5.2**).

Table 5.2. Potential markers for detecting the adulteration of GM with CM selected by VID procedure[1]

Adulterated Goat Milk		
Identity	Chemical class	VID
Phosphocholine	Amino acid	0.92
Orotate	Carboxylic acid	0.73
Glutamate	Amino acid	-0.77

[1]Only compounds with an absolute VID coefficients value of higher that 0.700 are selected. Compounds are listed in decreasing order of VID coefficient. The positive or negative VID coefficients indicate an increase or a decrease as a function of adulteration.

Based on **Table 5.2**, there are three most contributing variables in the class discrimination in the PLS-R model. The high VID means the compound is increasing as a function of the adulteration concentration *vice versa*. In another words, the amount of phosphocholine and orotate have increase due to the adulteration, while glutamate seems to decrease due to the adulteration.

The result of the current chapter matches the list of the discriminant markers used for distinguishing CM from GM and SM in **Chapter 4**, in which, orotate, phosphocholine, and glutamate were selected (see **Table 4.2**). In both chapter 4 and the current result, orotate and phosphocholine had positive VID value, while glutamate had negative VID value.

Boxplots were made using R-studio to depict the relative quantification of the discriminating compounds, in which Tukey's range test was performed. Boxplots depicted by the same letter indicates that there is no significant difference in the amounts of metabolites present in milk samples. On the other hand, different letters indicate the presence of significant difference.

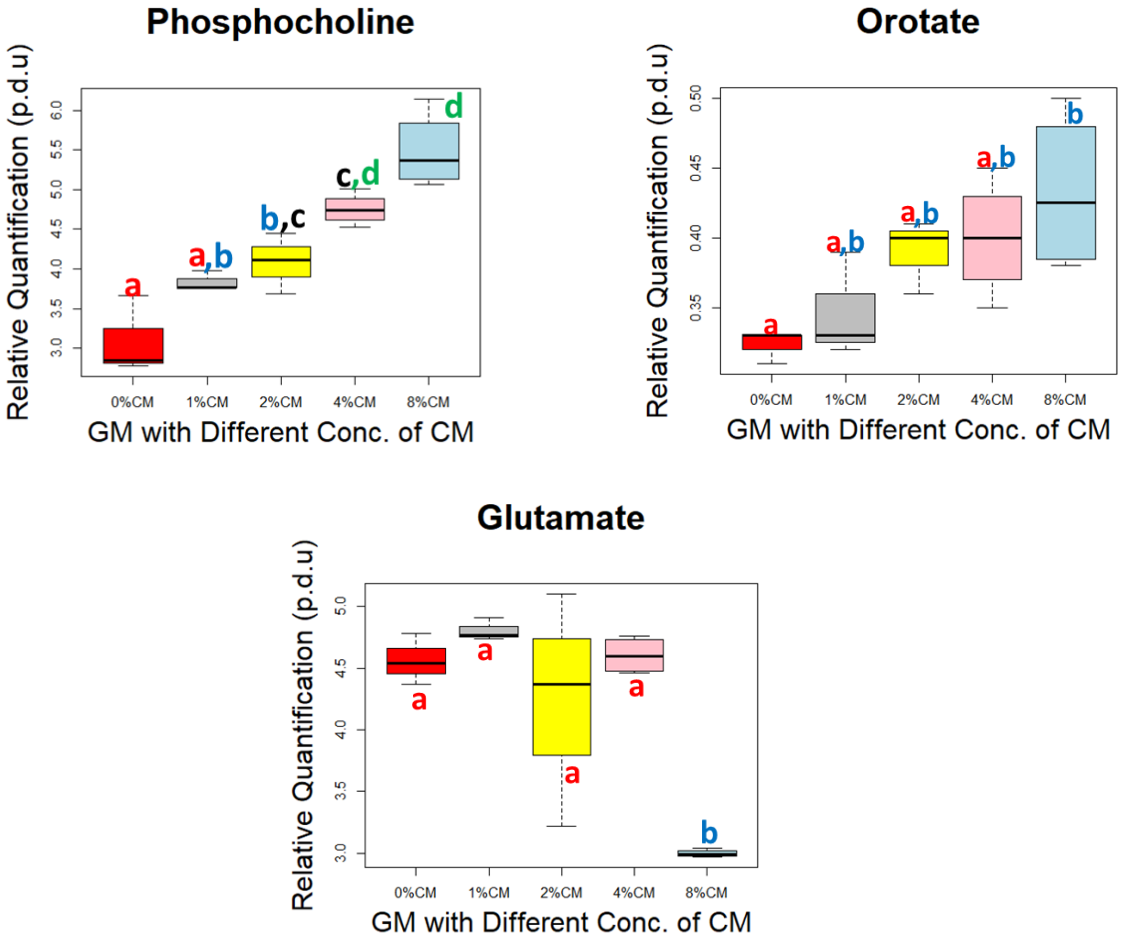

Figure 5.5. Discriminant compounds for comparison of unadulterated GM and adulterated GM with different concentration of CM (1%, 2%, 4%, and 8%), selected through VID procedure[1]

[1]The unit used for the relative quantification of each metabolites was p.d.u (procedure defined unit), in which it was based on the value of the internal standard (TSP) used in NMR spectrum.

As can be seen on **Figure 5.5**, the levels of phosphocholine and orotate were the lowest in the unadulterated GM and seems to increase as a function of the level of adulteration. Moreover, the level of phosphocholine and orotate increases in a linear fashion as the adulterant concentration increases. On the contrary, the level of glutamate was the lowest in GM adulterated with 8% of CM and the highest in GM adulterated with 1% of CM. Therefore, it was unclear whether the glutamate decreases as CM increases. When Tukey's test is taken into

consideration, phosphocholine and orotate were the interesting ones to be selected as potential markers of adulterations with CM.

The finding of the current study was supported by the previous study by Klein et al. (2012), in which phosphocholine was selected to be a biomarker for CM. Apart from that, phosphocholine is also useful as an indication of ketosis in dairy cows. In another study, phosphocholine was found to be an indicator of early lactation. Phosphocholine concentration decreases as the lactation stage progresses (Artegoitia et al., 2014).

As milk metabolites that is not well studied, there was limited information regarding orotate. Even though, orotate was found as milk components in various CM metabolomics studies, it is unknown whether it is a good discriminant marker of CM (Scano, Cusano, Caboni, & Consonni, 2019; Sundekilde, Larsen, et al., 2013). Nevertheless, orotate present in CM is said to be an indicator of metabolic deficiency of uridine monophosphate synthase (Zaalberg, Buitenhuis, Sundekilde, Poulsen, & Bovenhuis, 2020).

Overall, there was no information to confirm orotate as a good discriminant marker of CM. Additionally, based on the Tukey's test, the means for orotate was classified into two different groups, while the means for phosphocholine was classified into four groups. Thus, based on these findings, only phosphocholine can be proposed as a potential marker of adulteration in NZ GM with different concentrations of CM.

5.4.2 Adulteration of NZ Sheep Milk (SM) with Different Concentration of Cow Milk (CM)

The result for the detection of adulteration of NZ SM with CM is explained in the following sections:

5.4.2.1 Result from 1H- NMR Spectra of Adulterated SM

The representative NMR spectra coming from NZ SM adulterated with different concentration of CM is shown on **Figure 5.6**.

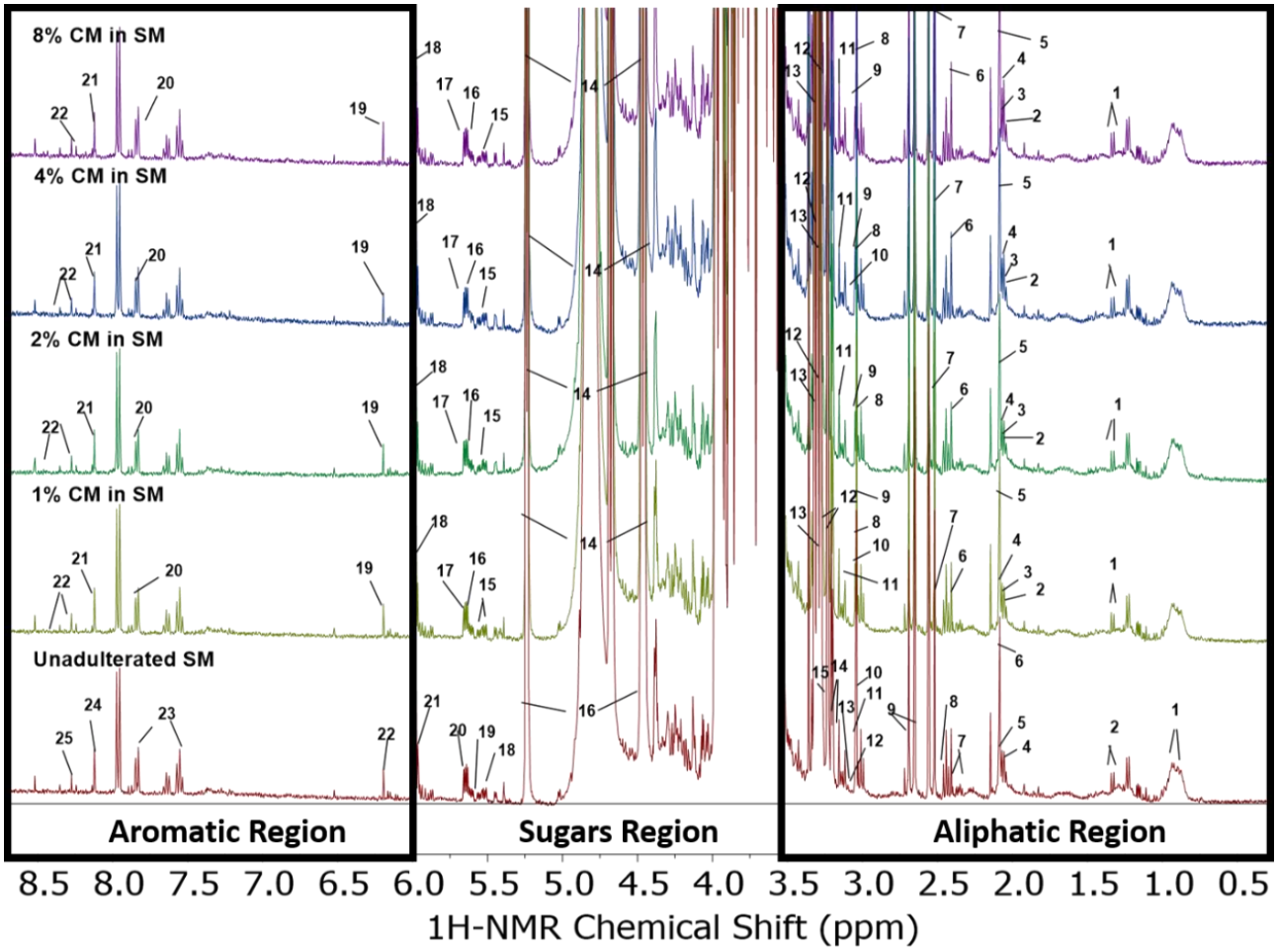

Figure 5.6. NMR Spectra of unadulterated SM with SM adulterated with different concentration of CM (1%, 2%, 4%, and 8%)

In line with the observation in the previous sections, the NMR spectra can be divided into three main regions, aliphatic, sugar, and aromatic region. Similar to the GM, only metabolites discriminating the different milk types (in Chapter 4) were considered for the adulteration study. In total, there were 7 carbohydrates, 6 amino acids, 6 carboxylic acids, 1 nucleotide, 1 nucleobase, and 1 nucleosides considered for the adulteration study. More details regarding the identity, chemical shifts, multiplicity, assignment, and the type of compounds found in NMR spectra of adulterated SM can be seen in **Table 5.3**. On the other hand, the details regarding

the compounds found in unadulterated SM can be seen in **Table 4.1**. Subsequently, advanced chemometrics were performed to study the changes in metabolites profile of SM as a result of adulteration.

Table 5.3. Metabolites assignment from 1H-NMR spectra of NZ SM adulterated with different concentration of CM (1%, 2%, 4%, and 8%)

Peak	Identity	δ_{1H} (multiplicity)[1]	Assignment	Type of Compounds	Presence	Ref[2]
1	Lactate	1.33(d)	CH_3	Carboxylic acid	All Milk	[1],[5]
2	N-acetyl glucosamine	2.04(s)	CH_3	Carbohydrate	All Milk	[6]
3	N-acetyl carbohydrate1	2.06(s)	CH_3	Carbohydrate	All Milk	[3]
4	N-acetyl carbohydrate2	2.07(s)	CH_3	Carbohydrate	All Milk	[3]
5	Glutamate	2.09(d)	β-CH_2	Amino Acid	All Milk	[3]
6	Succinic acid	2.41(s)	CH_2	Carboxylic acid	All Milk	[1]
7	Citrate	2.54(d)	CH_2	Tricarboxylic acid	All Milk	[4]
8	Creatine	3.03(s)	CH_3	Amino Acid	1%,2%,4% CM/SM	[2]
9	Phosphocreatine	3.04(s)	CH_3	Amino Acid	All Milk	[1],[6]
10	Creatinine	3.05(s)	CH_3	Amino Acid	1%,2%,8% CM/SM	[5]
11	Malonic acid	3.12(s)	CH_2	Dicarboxylic acid	All Milk	[2],[3]
12	Phosphocholine	3.20(m)	CH_3	Amino Acid	All Milk	[1]
13	Carnitine	3.23(s)	CH_3	Amino Acid	1%, 2% CM/SM	[7]
14	Lactose	4.68(d)	CH	Carbohydrate	All Milk	[1],[3]
		5.24(d)	CH	Carbohydrate	All Milk	[1],[4]
15	Glucose-1-phosphate	5.56(m)	CH	Carbohydrate	All Milk	[6],[7]
16	UDP glucoronate	5.61(dd)	CH	Carbohydrate	All Milk	[2]
17	UDP galactose	5.65(dd)	CH	Carbohydrate	All Milk	[1],[7]
18	Guanosine monophosphate	5.97(d)	CH	Nucleotides	All Milk	[1],[7]
19	Orotate	6.19(s)	CH	Carboxylic acid	All Milk	[2],[3],[6]
20	Hippuric acid	7.83(m)	CH_2	Carboxylic acid	All Milk	[1]
21	Adenine	8.12(s)	CH	Nucleobases	All Milk	[2],[3]
22	Inosine	8.24(s)	CH	Nucleosides	All Milk	[5],[6]

[1]Multiplicity: s=singlet; d=doublet; t=triplet; dd=double of doublets; m=multiplets.

[2]References: [1]=Human Metabolome Database (HMDB; http://hmdb.ca/); [2]=(Klein et al., 2010); [3]=(Sundekilde, Larsen, et al., 2013), [4]= (Sundekilde, Poulsen, et al., 2013), [5]=(Li, Yu, et al., 2017),[6]=(Zhao et al., 2017),[7]=(Sanchez et al., 2020)

5.4.2.2 Detecting the Adulteration using Chemometrics with Principal Component Analysis (PCA)

Unsupervised PCA analysis was performed (result not shown) to explore the trend and patterns due to the systematic adulteration of SM with different amount of CM and to check for outliers. PCA was useful in visualising the data trend as a function of adulteration. For further investigation, PLS-R is applied to explain the changes in the metabolites in SM as the adulterant concentrations increases.

5.4.2.3 Detecting the Adulteration using Chemometrics with Partial Least Square Regression Analysis (PLS-R)

Following PCA analysis, PLS-R is also used to identify discriminant compounds that can be a candidate marker of adulteration. Thus, a leave one out cross validation was performed. Cumulative variance and RMSECV is shown in ***Figure 5.7*** and ***Figure 5.8*** below.

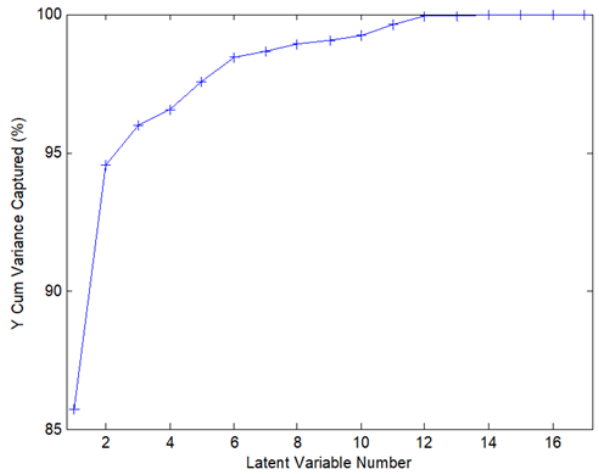

Figure 5.7. Cumulative Variance Graph - SM adulterated with CM (1%, 2%, 4%, and 8%) (PLSR)

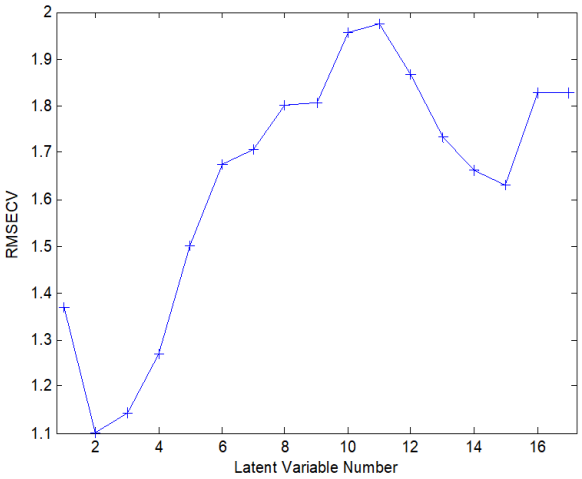

Figure 5.8. Root Mean Square Error of Cross Validation Graph - SM adulterated with CM (1%, 2%, 4%, and 8%) (PLS-R)

Based on the cross validation, 2 LVs were chosen for the PLS-R model. The 2 LVs represent 55.99% and 94.55% of the cumulative variance for X and Y variables, respectively. After cross validation, the first two LVs explaining the highest variance in the data were used to construct score plot and loading plots to illustrate the interrelationship in the data. The 2 chosen LVs were used to construct score plot and loadings plot, and these can be seen in appendix (**Figure A.3** and **Figure A.4**).

When scores plot and loadings plot were combined, bi-plot were generated. In PLS-R, a bi-plot typically portrays the regression relationship between a set of predictors and the response variables (Y) (Oyedele & Lubbe, 2015).

In the present study, the bi-plot (**Figure 5.9**) displays the interrelationships between the milk metabolites with the level of adulterations in milk samples. The bi-plots shows a clear relationship between the milk samples and the adulterants concentration, in which the loadings position of milk metabolites changed in accordance with the increasing level of adulterants concentration. There is a trend moving from left to the right side in accordance with the increase of adulteration concentrations. The unadulterated SM is located on the far left, followed by SM adulterated with 1% of CM, SM adulterated with 2% of CM, SM adulterated with 4% of CM, and lastly SM adulterated with 8% of CM. This shows a clear change in the metabolite profile of SM due to adulteration with CM.

As mentioned in **Chapter 4**, variables whose loadings are co-located away from the centre of the coordinate may be inferred to be correlated. Moreover, variables with loadings in a given position on the plot contribute heavily to observations whose scores are found in a similar position in a scores plot (Worley & Powers, 2013). Thus, the loadings strongly correlated with scores for each milk types identify metabolites that may uniquely describe or characterize the selected adulterated sample.

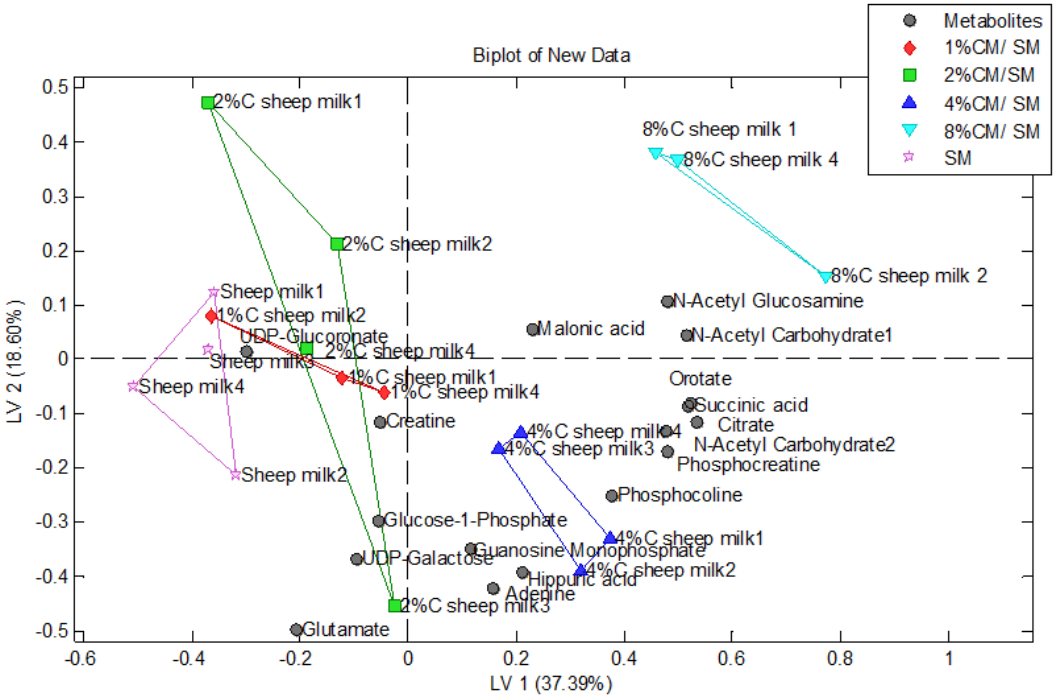

Figure 5.9. PLS-R Bi-plot for detecting SM adulteration with different concentration of CM (1%, 2%, 4%, and 8%)

To conclude, the established PLS-R model was successful in classifying the adulterated samples from the control, even at a very small concentration of CM. To accurately identify these candidate markers of adulteration, VID coefficients were calculated (see next section).

5.4.2.4 Discriminant Markers Selection of Adulterated Sheep Milk (SM)

The VID procedure enabled the selection of 8 discriminant compounds. All the discriminant compounds present in the adulterated SM samples are listed in decreasing order of VID coefficient, as shown in **Table 5.4**.

Table 5.4. Potential markers for detecting the adulteration of SM with CM selected by VID procedure[1]

Adulterated Sheep Milk		
Identity	Chemical class	VID
N-Acetyl Carbohydrate1	Carbohydrate	0.87
Citrate	Tricarboxylic acid	0.84
Orotate	Carboxylic acid	0.84
N-Acetyl Glucosamine	Carbohydrate	0.83
Succinic Acid	Dicarboxylic acid	0.83
N-Acetyl Carbohydrate2	Carbohydrate	0.74
Phosphocreatine	Amino acid	0.73

[1]Only compounds with an absolute VID coefficients value of higher that 0.700 are selected. Compounds are listed in decreasing order of VID coefficient. The positive or negative VID coefficients indicate an increase or a decrease as a function of adulteration

Table 5.4 shows, the most contributory variable in the class discrimination in the PLS-R model including N-acetyl carbohydrate1, citrate, orotate, N-acetyl glucosamine, succinic acid, N-acetyl carbohydrate2, and phosphocreatine.

As the SM used in the present chapter has been adulterated with CM, the result of the current chapter matches the list of the discriminant markers used for distinguishing CM from GM and SM in **Chapter 4**, where orotate, N-acetyl glucosamine, N-acetyl carbohydrate1, N-acetyl carbohydrate2, and citrate were all selected because of their high positive VID coefficients (see **Table 4.2**). Nevertheless, succinic acid in CM had a negative VID coefficient, while phosphocreatine was not selected as discriminant marker for CM.

Following VID procedures, boxplots were constructed to depict the relative quantification of the selected markers at different levels of adulteration in the SM samples. Furthermore, Tukey's test of significance was performed, where a similar letter indicates that there is no significant difference between the samples, vice versa.

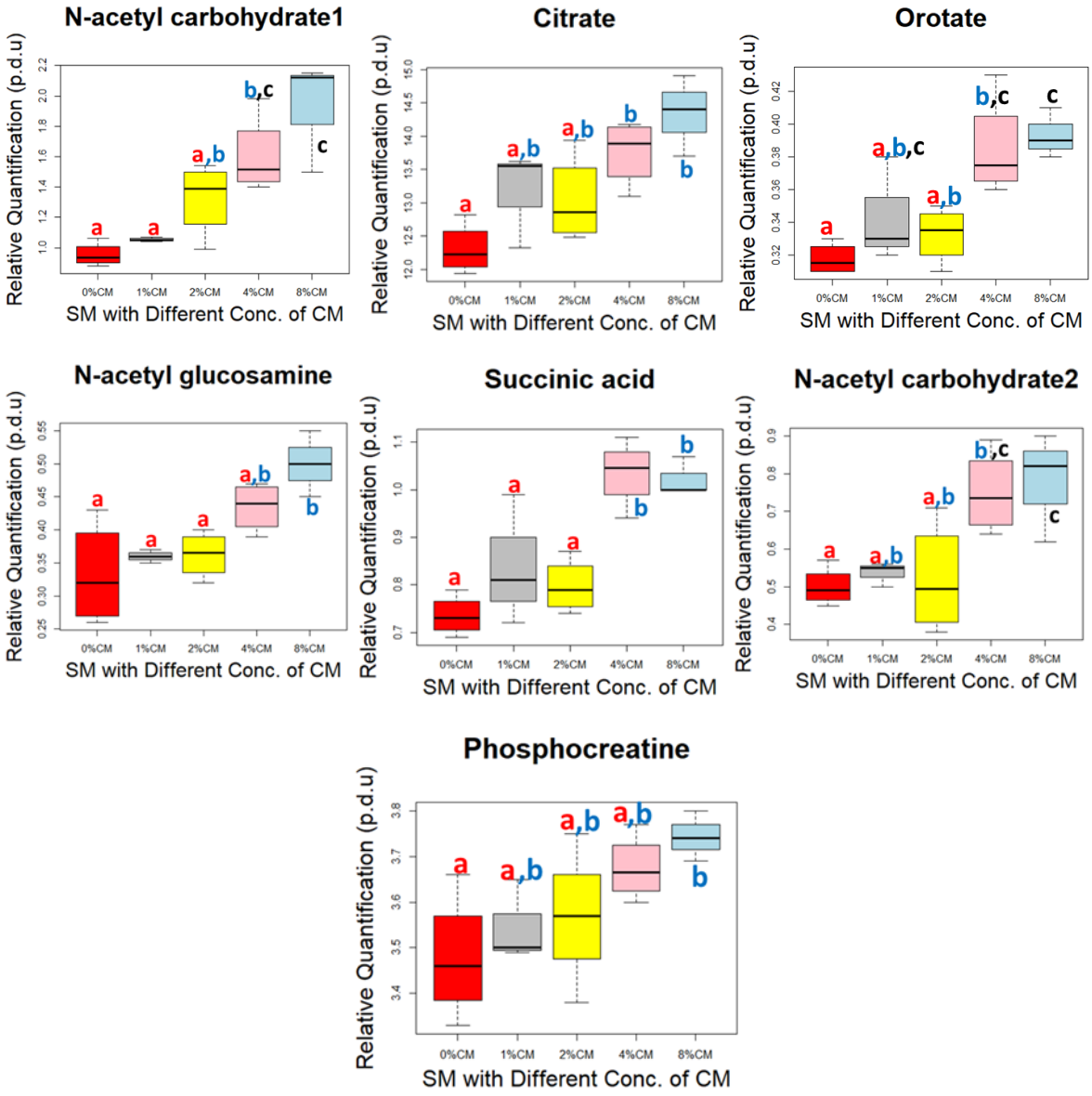

Figure 5.10. Discriminant compounds for comparison of unadulterated SM and adulterated SM with different concentration of CM (1%, 2%, 4%, and 8%), selected through VID procedure[1]

[1]The unit used for the relative quantification of each metabolites was p.d.u (procedure defined unit), in which it was based on the value of the internal standard (TSP) used in NMR spectrum.

As seen on **Figure 5.10**, the amount of N-acetyl carbohydrate1, citrate, orotate, N-acetyl glucosamine, succinic acid, N-acetyl carbohydrate2, and phosphocreatine increases as the adulteration concentration increases.

The result of the present study matches the findings from the literature. In the NMR study by Hu et al. (2004), citrate and N-acetyl carbohydrates were both found in CM. In another study, all 7 compounds were found as metabolite constituent in CM (Sundekilde, Larsen, et al., 2013,). N-acetyl carbohydrates, citrate, orotate, and phosphocreatine were also present in the

integrated metabolomics study of milk from heat-stressed dairy cows (Tian et al., 2016). Yang et al. (2016) had previously identified succinic acid as metabolic biomarkers to differentiate CM from other ruminant milk.

Based on our literature research, there are no studies investigating the characterization and detection of adulteration in SM using ^{1}H-NMR. Most studies involving NMR-spectroscopy on SM were concentrated on the detection of adulteration in cheeses originated from SM, e.g., imitation cheeses (Monakhova, Godelmann, Andlauer, Kuballa, & Lachenmeier, 2013) and Fossa cheese (Scano, Cagliani, & Consonni, 2019). Besides, there has only been one untargeted metabolomics study of SM obtained from different grazing system. From the study, compounds including succinic acid, inosine, hippuric acid were deemed as important in SM (Scano, Carta, Ibba, Manis, & Caboni, 2020). Therefore, it was a challenge to link the result observed in the present work with the existing literature. This demonstrated the research gap and need for further study in this research area.

Based on the Tukey's test results, N-acetyl carbohydrates and orotate can be proposed as potential markers of adulteration. There were significant differences in the presence of N-acetyl carbohydrates and orotate in between the adulterated milk types in relation to the increasing adulterant concentration. On the other hand, citrate, N-acetyl glucosamine, succinic acid, and phosphocreatine did not have significant differences. Hence, N-acetyl carbohydrate and orotate were selected as candidate markers for detecting the adulteration of SM with CM.

5.4.3 Summary of Chapter 5

NMR-based metabolomics technique was coupled with advanced chemometrics (unsupervised PCA and supervised PLS-R) was used to detect adulteration of NZ GM and SM with different levels of CM. It was also attempted to identify potential markers to detect adulteration of these high value dairy products.

Based on VID calculation, 3 metabolites including orotate, phosphocholine, and glutamate were selected as discriminant compounds for GM. On the other hand, 7 metabolites including N-acetyl carbohydrate1, citrate, orotate, N-acetyl glucosamine, succinic acid, N-acetyl carbohydrate2, and phosphocreatine were selected discriminating the different levels of adulteration in SM. To ensure their capability as discriminant markers in the adulterated samples, Tukey's test was performed on the selected compounds. Overall, phosphocholine can be proposed as a candidate marker for detecting adulteration of GM with CM, while N-acetyl carbohydrate and orotate were chosen as potential markers of adulteration of SM with CM.

Chapter 6 General Discussion, Conclusion, Limitations, and Future Outlook

The general discussion, conclusion, limitations, and the future outlook of the current MSc research can be found as follows.

6.1 General Discussion

The potential application of NMR-based metabolomics on detection of adulterations in different types of milk has been mentioned in the literature. As a powerful analytical technique, NMR allows detailed investigation of quantitative and qualitative analysis of milk samples. The resulting data from NMR-based metabolomics technique provides a wealth of information about the samples. As a result, it is important to apply a comprehensive data analysis technique that allows data explorations. One such method is chemometrics. Chemometrics is heavily used in the field of analytical chemistry and metabolomics. It is a useful tool that can provide unsupervised and supervised approaches for exploratory analysis, regression analysis, and data classifications and marker selection.

In **Chapter 4,** the metabolite profile of NZ milk powders from cow, goat, and sheep were successfully characterised by NMR-based metabolomics combined with chemometrics. NMR fingerprinting technique effectively identified 17, 24, and 23 metabolites present in the liquid fractions of CM, GM, and SM, respectively. The data obtained from NMR was processed and interpreted using unsupervised (PCA) and supervised (PLS-DA) chemometrics approach.

The capability of PCA as a dimensionality reduction technique and PLS-DA as a classification approach were clearly demonstrated in **Chapter 4**, in which CM, GM, and SM were found to be different from each other. Specifically, CM has different metabolite composition, while GM and SM appears to have comparable metabolite composition. The results from PLS-DA, were then used to calculate VID coefficient to identify discriminant compounds for each milk. From the result, GM was found to have the highest number of discriminant compounds, followed by CM, and SM with 9, 8, and 5 respectively. Altogether, when further verification techniques are applied, the current result will be a great addition into the compositional databases of NZ CM, GM, and SM. On top of that, by characterising metabolites present in NZ CM, GM, and SM it is now feasible to differentiate among the three milk types. To conclude, Chapter 4 resulted in

the selection of discriminant compounds classifying the three milk types, which was then used as a basis for the adulteration experiment in the following chapter.

In **Chapter 5,** NMR-fingerprinting technique coupled with advanced chemometrics was applied to detect and identify candidate markers of adulteration in NZ GM and SM. The SM and GM samples were systematically adulterated with different concentration of CM (1%, 2%, 4%, and 8%). For both GM and SM samples, the established PLS-R model was successful in classifying the adulterated samples from the control, even at a very small concentration of CM Following this, VID coefficients were calculated, and Tukey's test was performed in order to identify candidate adulteration markers for each of the high-value NZ dairy products. Phosphocholine can be selected as markers of adulteration of GM with CM, while N-acetyl carbohydrates and orotate can be selected as potential markers of adulteration of SM with CM.

6.2 Conclusion

Overall, the current study has successfully addressed the research gaps. NMR-based metabolomics coupled with advanced chemometrics were effectively implemented to identify and characterize the metabolites present in the unadulterated and adulterated NZ milk samples. The metabolites responsible for distinguishing between unadulterated and adulterated milk samples were defined. In the future, these metabolites could be used as potential markers to optimize, monitor, and predict quality parameters of the different milk samples during manufacturing processes to protect them from adulteration.

6.3 Limitations of the current study and Future Outlook

Based on the result of the current study, NMR-based metabolomics technique combined with advanced chemometrics is suitable to characterise and detect markers of adulteration in NZ high-value dairy products. However, the current study has limitations that needs to be addressed for the future work. The following suggestions are mentioned below:

1. The current study is a novel study involving the characterisation and detection of adulteration in NZ dairy products. In the present work, the milk samples were sourced from a single supplier. CM sample was obtained from a local supermarket in Dunedin, GM from NIG Nutritionals Limited, and lastly, SM from Blue River Dairy. Additionally, samples from different sources may naturally have variations, and any processing steps pursued on the samples may further contribute to differences in metabolite profiles. Therefore, in future studies, milks from different suppliers must be considered to validate the observations.

2. In the present study, only water-soluble fraction of the milk samples was considered. This, therefore, affected the number of metabolites that was identified in the present study. In the future, other fractions such as the lipid-fraction can be considered to correlate with the metabolites detected in the present work.

3. Next, the other factor relating to the milk samples was not considered. As the milk powder came from a single supplier, factors like environment, lactation stage, seasonal variation, breed, quality of feed, and weight of the animals were not taken into account. These factors must be studied as they have a significant influence on milk metabolites compositions and its nutritional quality.

4. It was challenging to identify metabolites present in the milk samples, as there was a scarcity of information regarding the metabolite composition of NZ GM and SM. In the future, there is a need to build a robust database.

5. In this work, the NMR data was processed using the linear chemometrics method (PCA, PLS-DA, and PLS-R). However, there is a recent interest in the usage of non-linear chemometrics method such as artificial neural networks (ANNs) and decision tree to solving the problems of exploratory analysis, prediction, and classification. As a result,

it might be worth considering applying non-linear chemometrics approach in future studies (especially when a data obtained from multiplatform approach is considered).

6. The use of other metabolomic fingerprinting techniques to further characterise the metabolites present in the different milk type was beyond the scope of this work. Therefore, in future studies, techniques such as LC-Q-TOF-MS and GC-MS, or other infrared-based methods (FTIR, NIR) could be used to characterize and detect adulteration of high value dairy products, and the data compared with the findings of this study.

References

Abd El-Salam, M. H., & El-Shibiny, S. (2011). A comprehensive review on the composition and properties of buffalo milk. *Dairy Science & Technology, 91*(6), 663. doi:10.1007/s13594-011-0029-2

Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (PLS Regression). *WIREs Computational Statistics, 2*(1), 97-106. doi:10.1002/wics.51

Abu-Aboud, O., & Weiss, R. (2012). New Opportunities from the cancer metabolome. *Clinical chemistry, 59*. doi:10.1373/clinchem.2012.184598

Agostoni, C., & Turck, D. (2011). Is cow's milk harmful to a child's health? *Journal of Pediatric Gastroenterology and Nutrition, 53*(6), 594-600. doi:10.1097/MPG.0b013e318235b23e

Agudiez, M., Martinez, P. J., Martin-Lorenzo, M., Heredero, A., Santiago-Hernandez, A., Molero, D.,Alvarez-Llamas, G. (2020). Analysis of urinary exosomal metabolites identifies cardiovascular risk signatures with added value to urine analysis. *BMC Biology, 18*(1), 192. doi:10.1186/s12915-020-00924-y

Ahirwar, R., Harilal, P., Srihari, K., & Pandey, M. (2015). Quality changes in milk adulterated with detergent, urea, ammonium sulphate and neutralizers. *Asian Journal of Dairy and Food Research, 34*(4), 285-289.

Aiello, A., Pizzolongo, F., Manzo, N., & Romano, R. (2019). A new method to distinguish the milk adulteration with neutralizers by detection of lactic acid. *Food Analytical Methods, 12*(11), 2555-2561. doi:10.1007/s12161-019-01594-5

Akalin, A., & Gonc, S. (1996). The determination of orotic acid in ruminant milks using HPLC.

Alhussien, M. N., & Dang, A. K. (2018). Milk somatic cells, factors influencing their release, future prospects, and practical utility in dairy animals: An overview. *Veterinary World, 11*(5), 562-577. doi:10.14202/vetworld.2018.562-577

Alichanidis, E., Moatsou, G., & Polychroniadou, A. (2016). Chapter 5 - Composition and properties of non-cow milk and products. In E. Tsakalidou & K. Papadimitriou (Eds.), *Non-Bovine Milk and Milk Products* (pp. 81-116). San Diego: Academic Press.

Alonso, A., Marsal, S., & Julià, A. (2015). Analytical methods in untargeted metabolomics: State of the Art in 2015. *Frontiers in Bioengineering and Biotechnology, 3*(23). doi:10.3389/fbioe.2015.00023

Ambrose, A., & Cho, B.-K. (2014). A review of technologies for detection and measurement of adulterants in cereals and cereal products. *Journal of Biosystems Engineering, 39*(4), 357-365.

Amha, Y. T. N. (2015). Use of different non protein nitrogen sources in ruminant nutrition: A review. *Advances in Life Science and Technology, 29*, 100-105.

Amills, M., Capote, J., & Tosser-Klopp, G. (2017). Goat domestication and breeding: A jigsaw of historical, biological and molecular data with missing pieces. *Anim Genet, 48*(6), 631-644. doi:10.1111/age.12598

Ander, B. P., Dupasquier, C. M., Prociuk, M. A., & Pierce, G. N. (2003). Polyunsaturated fatty acids and their effects on cardiovascular disease. *Experimental and Clinical Cardiology, 8*(4), 164-172. Retrieved from https://pubmed.ncbi.nlm.nih.gov/19649216

Andreotti, G., Trivellone, E., & Motta, A. (2006). Characterization of buffalo milk by 31P-nuclear magnetic resonance spectroscopy. *Journal of Food Composition and Analysis, 19*(8), 843-849.

Angelopoulou, I., Botsialas, A., Salapatas, A., Petrou, P. S., Haasnoot, W., Makarona, E., . . . Kakabakos, S. E. (2015). Assessment of goat milk adulteration with a label-free monolithically integrated optoelectronic biosensor. *Analytical and Bioanalytical Chemistry, 407*(14), 3995-4004. doi:10.1007/s00216-015-8596-3

Arcena, M. R., Kebede, B., Leong, S. Y., Silcock, P., & Oey, I. (2020). Feasibility of using integrated fingerprinting, profiling and chemometrics approach to understand (bio) chemical changes throughout commercial red winemaking: A case study on Merlot. *Food Research International, 127*, 108767. doi:https://doi.org/10.1016/j.foodres.2019.108767

Artegoitia, V. M., Middleton, J. L., Harte, F. M., Campagna, S. R., & de Veth, M. J. (2014). Choline and choline metabolite patterns and associations in blood and milk during lactation in dairy cows. *PloS One, 9*(8), e103412-e103412. doi:10.1371/journal.pone.0103412

Atkins, P. J. (1991). Sophistication detected: Or, the adulteration of the milk supply, 1850–1914. *Social History, 16*(3), 317-339. doi:10.1080/03071029108567811

Ayza, A., & Belete, E. (2015). Food adulteration: its challenges and impacts. *Food Sci Qual Manag, 41*, 50-56.

Azad, T., & Ahmed, S. (2016). Common milk adulteration and their detection techniques. *International Journal of Food Contamination, 3*. doi:10.1186/s40550-016-0045-3

Ball, G. F. (2008). *Vitamins: their role in the human body.*USA: John Wiley & Sons.

Ballabio, D., & Consonni, V. (2013). Classification tools in chemistry. Part 1: linear models. PLS-DA. *Analytical Methods, 5*(16), 3790-3798. doi:10.1039/C3AY40582F

Ballingall, J., & Pambudi, D. (2017). Dairy trade's economic contribution to New Zealand. *New Zealand Institute of Economic Research.*

Balthazar, C. F., Conte Júnior, C. A., Moraes, J., Costa, M. P., Raices, R. S. L., Franco, R. M.,Silva, A. C. O. (2016). Physicochemical evaluation of sheep milk yogurts containing different levels of inulin. *Journal of Dairy Science, 99*(6), 4160-4168. doi:https://doi.org/10.3168/jds.2015-10072

Balthazar, C. F., Pimentel, T. C., Ferrão, L. L., Almada, C. N., Santillo, A., Albenzio, M.,Cruz, A. G. (2017). Sheep milk: Physicochemical characteristics and relevance for functional food Development. *Comprehensive Reviews in Food Science and Food Safety, 16*(2), 247-262. doi:10.1111/1541-4337.12250

Bansal, S., Singh, A., Mangal, M., Mangal, A. K., & Kumar, S. (2017). Food adulteration: Sources, health risks, and detection methods. *Critical Reviews in Food Science and Nutrition, 57*(6), 1174-1189. doi:10.1080/10408398.2014.967834

Barboza, D. (2009). Death sentences given in Chinese milk scandal. *The New York Times*. Retrieved from https://www.nytimes.com/2009/01/22/world/asia/22iht-milk.1.19587787.html

Barham, G. S., Khaskheli, M., Soomro, A. H., & Nizamani, Z. A. (2014). Extent of extraneous water and detection of various adulterants in market milk at Mirpurkhas, Pakistan. *J Agri Vet Sci, 7*(3), 83-89.

Barłowska, J., Szwajkowska, M., Litwińczuk, Z., & Król, J. (2011). Nutritional Value and Technological Suitability of Milk from Various Animal Species Used for Dairy Production. *Comprehensive Reviews in Food Science and Food Safety, 10*(6), 291-302. doi:10.1111/j.1541-4337.2011.00163.x

Bencini, R., & Pulina, G. (1997). The quality of sheep milk: A review. *Australian Journal of Experimental Agriculture, 37*, 485-504. doi:10.1071/EA96014

Bernard, J. K., & Tao, S. (2019). Feeding for improved yield of milk components. *UGA Cooperative Extension Bulletin*, (1512). Retrieved from https://secure.caes.uga.edu/extension/publications/files/pdf/B%201512_1.PDF

Berrueta, L. A., Alonso-Salces, R. M., & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A, 1158*(1), 196-214. doi:https://doi.org/10.1016/j.chroma.2007.05.024

Biancolillo, A., & Marini, F. (2018). Chemometric methods for spectroscopy-based pharmaceutical analysis. *Frontiers in Chemistry, 6*(576). doi:10.3389/fchem.2018.00576

Bilandžić, N., Božić, Đ., Đokić, M., Sedak, M., Solomun Kolanovic, B., Varenina, I., & Cvetnić, Ž. (2014). Assessment of aflatoxin M1 contamination in the milk of four dairy species in Croatia. *Food Control, 43*, 18–21. doi:10.1016/j.foodcont.2014.02.044

Bobková, A., Židek, R., Flimelová, E., Bobko, M., & Fiková, M. (2009). Application of PCR method for milk adulteration and milk products identification. *Potravinarstvo, 3*, 1-3.

Borková, M., & Snášelová, J. (2005). Possibilities of different animal milk detection in milk and dairy products - A review. *Czech Journal of Food Sciences, 23*, 41-50. doi:10.17221/3371-CJFS

Brantsæter, A. L., Olafsdottir, A. S., Forsum, E., Olsen, S. F., & Thorsdottir, I. (2012). Does milk and dairy consumption during pregnancy influence fetal growth and infant birthweight? A systematic literature review. *Food & Nutrition Research, 56*, 10.3402/fnr.v3456i3400.20050. doi:10.3402/fnr.v56i0.20050

Brown, J. R., Law, A. J., & Knight, C. H. (1995). Changes in casein composition of goats' milk during the course of lactation: Physiological inferences and technological implications. *Journal of Dairy Research, 62*(3), 431-439.

Brown, S., Tauler, R., & Walczak, B. (2020). *Comprehensive chemometrics: chemical and biochemical data analysis.*Spain: Elsevier.

Burrow, K., Young, W., McConnell, M., Carne, A., & Bekhit, A. E.-D. (2018). Do dairy minerals have a positive effect on bone health? *Comprehensive Reviews in Food Science and Food Safety, 17*(4), 989-1005. doi:10.1111/1541-4337.12364

Bus, A. E. M., & Worsley, A. (2003). Consumers' sensory and nutritional perceptions of three types of milk. *Public Health Nutrition, 6*(2), 201-208. doi:10.1079/PHN2002417

Butler, G., Stergiadis, S., Seal, C., Eyre, M., & Leifert, C. (2011). Fat composition of organic and conventional retail milk in northeast England. *Journal of Dairy Science, 94*(1), 24-36. doi:10.3168/jds.2010-3331

Caboni, P., Murgia, A., Porcu, A., Demuru, M., Pulina, G., & Nudda, A. (2016). Gas chromatography-mass spectrometry metabolomics of goat milk with different polymorphism at the S1-casein genotype locus. *Journal of Dairy Science, 99*(8), 6046-6051. doi:10.3168/jds.2015-10537

Callaway, J. C. (2004). Hempseed as a nutritional resource: An overview. *Euphytica, 140*(1), 65-72. doi:10.1007/s10681-004-4811-6

Camacho, J., & Ferrer, A. (2014). Cross-validation in PCA models with the element-wise k-fold (ekf) algorithm: Practical aspects. *Chemometrics and Intelligent Laboratory Systems, 131*, 37-50. doi:https://doi.org/10.1016/j.chemolab.2013.12.003

Cao, G., Song, Z., Hong, Y., Yang, Z., Song, Y., Chen, Z.,Cai, Z. (2020). Large-scale targeted metabolomics method for metabolite profiling of human samples. *Analytica Chimica Acta, 1125*, 144-151. doi:https://doi.org/10.1016/j.aca.2020.05.053

Capuano, E., Rademaker, J., van den Bijgaart, H., & van Ruth, S. M. (2014). Verification of fresh grass feeding, pasture grazing and organic farming by FTIR spectroscopy analysis of bovine milk. *Food Research International, 60*, 59-65.

Caroli, A., Poli, A., Ricotta, D., Banfi, G., & Cocchi, D. (2011). Invited review: Dairy intake and bone health: A viewpoint from the state of the art. *Journal of Dairy Science, 94*(11), 5249-5262. doi:https://doi.org/10.3168/jds.2011-4578

Carpio, A., Bonilla-Valverde, D., Arce, C., Rodríguez-Estévez, V., Sánchez-Rodríguez, M., Arce, L., & Valcárcel, M. (2013). Evaluation of hippuric acid content in goat milk as a marker of feeding regimen. *Journal of Dairy Science, 96*(9), 5426-5434. doi:https://doi.org/10.3168/jds.2012-6396

Carpio, A., Rodríguez-Estévez, V., Sánchez-Rodríguez, M., Arce, L., & Valcárcel, M. (2010). Differentiation of organic goat's milk based on its hippuric acid content as determined by capillary electrophoresis. *Electrophoresis, 31*, 2211-2217. doi:10.1002/elps.200900728

Carrascal, L. M., Galván, I., & Gordo, O. (2009). Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos, 118*(5), 681-690. doi:10.1111/j.1600-0706.2008.16881.x

Cartoni, G., Coccioli, F., Jasionowska, R., & Masci, M. (1999). Determination of cows' milk in goats' milk and cheese by capillary electrophoresis of the whey protein fractions. *Journal of Chromatography A, 846*(1-2), 135-141. doi:10.1016/s0021-9673(98)01032-2

Cevallos-Cevallos, J. M., Reyes-De-Corcuera, J. I., Etxeberria, E., Danyluk, M. D., & Rodrick, G. E. (2009). Metabolomic analysis in food science: a review. *Trends in Food Science & Technology, 20*(11-12), 557-566.

Chan, E. Y. Y., Griffiths, S. M., & Chan, C. W. (2008). Public-health risks of melamine in milk products. *Lancet (London, England), 372*(9648), 1444-1445. doi:10.1016/S0140-6736(08)61604-9

Cheah, W. L., & Fang, M. (2020). HPLC-based chemometric analysis for coffee adulteration. *Foods, 9*(7), 880.

Chen, S., Kong, H., Lu, X., Li, Y., Yin, P., Zeng, Z., & Xu, G. (2013). Pseudotargeted metabolomics method and its application in serum biomarker discovery for hepatocellular carcinoma based on ultra high-performance liquid chromatography/triple quadrupole mass spectrometry. *Analytical Chemistry, 85*(17), 8326-8333. doi:10.1021/ac4016787

Cifuentes, A. (2006). Recent advances in the application of capillary electromigration methods for food analysis. *Electrophoresis, 27*(1), 283-303.

Claeys, W. L., Verraes, C., Cardoen, S., De Block, J., Huyghebaert, A., Raes, K.,Herman, L. (2014). Consumption of raw or heated milk from different species: An evaluation of the nutritional and potential health benefits. *Food Control, 42*, 188-201. doi:https://doi.org/10.1016/j.foodcont.2014.01.045

Clark, S., & Mora García, M. B. (2017). A 100-year review: Advances in goat milk research. *Journal of Dairy Science, 100*(12), 10026-10044. doi:https://doi.org/10.3168/jds.2017-13287

Consonni, R., & Cagliani, L. (2010). Nuclear magnetic resonance and chemometrics to assess geographical origin and quality of traditional food products. In *Advances in Food and Nutrition Research* (Vol. 59, pp. 87-165).Italy: Elsevier.

Coonan, C. (2013). New Zealand dairy firm tries to contain Chinese milk scandal. *The Irish Times*. Retrieved from https://www.irishtimes.com/business/agribusiness-and-food/new-zealand-dairy-firm-tries-to-contain-chinese-milk-scandal-1.1485403

Cruz-Teno, C., Pérez-Martínez, P., Delgado-Lista, J., Yubero-Serrano, E. M., García-Ríos, A., Marín, C.,López-Miranda, J. (2012). Dietary fat modifies the postprandial inflammatory state in subjects with metabolic syndrome: the LIPGENE study. *Molecular Nutrition & Food Research, 56*(6), 854-865. doi:10.1002/mnfr.201200096

Cui, J., Zhu, D., Su, M., Tan, D., Zhang, X., Jia, M., & Chen, G. (2019). The combined use of 1H and 2D NMR-based metabolomics and chemometrics for non-targeted screening of biomarkers and identification of reconstituted milk. *Journal of the Science of Food and Agriculture, 99*(14), 6455-6461. doi:10.1002/jsfa.9924

Cui, X., Zhang, J., Cai, W., & Shao, X. (2017). Chemometric algorithms for analyzing high dimensional temperature dependent near infrared spectra. *Chemometrics and Intelligent Laboratory Systems, 170*, 109-117. doi:https://doi.org/10.1016/j.chemolab.2017.08.010

Cui, Y., Xu, B., Zhang, X., He, Y., Shao, Y., & Ding, M. (2018). Diagnostic and therapeutic profiles of serum bile acids in women with intrahepatic cholestasis of pregnancy-a pseudo-targeted metabolomics study. *Clinica Chimica Acta, 483*, 135-141.

da Costa, T. H. M., Haisma, H., Wells, J. C. K., Mander, A. P., Whitehead, R. G., & Bluck, L. J. C. (2010). How much human milk do infants consume? Data from 12 countries using a

standardized stable isotope methodology. *The Journal of Nutrition, 140*(12), 2227-2232. doi:10.3945/jn.110.123489

DairyNZ. (2019). *New Zealand Dairy Statistic 2018-2019*. Report from Dairy New Zealand

Dakeishi, M., Murata, K., & Grandjean, P. (2006). Long-term consequences of arsenic poisoning during infancy due to contaminated milk powder. *Environmental Health, 5*, 31-31. doi:10.1186/1476-069X-5-31

Dalgleish, D. (1982). Milk proteins. Chemistry and physics. *Food Proteins*, 155-178.

Das, S., Goswami, B., & Biswas, K. (2016). Milk adulteration and detection: A review. *Sensor Letters, 14*, 4-18. doi:10.1166/sl.2016.3580

Davoodi, H., Esmaeili, S., & Mortazavian, A. M. (2013). Effects of Milk and Milk Products Consumption on Cancer: A Review. *Comprehensive Reviews in Food Science and Food Safety, 12*(3), 249-264. doi:10.1111/1541-4337.12011

Davoodi, S. H., Shahbazi, R., Esmaeili, S., Sohrabvandi, S., Mortazavian, A., Jazayeri, S., & Taslimi, A. (2016). Health-Related Aspects of Milk Proteins. *Iranian journal of Pharmaceutical Research : IJPR, 15*(3), 573-591. Retrieved from https://pubmed.ncbi.nlm.nih.gov/27980594

DCANZ. (2020). About the NZ Dairy Industry. Report from DCANZ. Retrieved from https://www.dcanz.com/about-the-nz-dairy-industry/

Debnath, A., Banerjee, S., Rai, C., & Roy, A. (2015). Qualitative detection of adulterants in milk samples from Kolkata and its surburban areas. *International Journal of Research in Applied, Natural and Social Sciences, 3*, 81-88.

DePeters, E. J., & Ferguson, J. D. (1992). Nonprotein Nitrogen and Protein Distribution in the Milk of Cows. *Journal of Dairy Science, 75*(11), 3192-3209. doi:https://doi.org/10.3168/jds.S0022-0302(92)78085-0

Dettmer, K., Aronov, P. A., & Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass spectrometry reviews, 26*(1), 51-78.

Diana, G., & Tommasi, C. (2002). Cross-validation methods in principal component analysis: A comparison. *Statistical Methods and Applications, 11*(1), 71-82. doi:10.1007/BF02511446

DiNicolantonio, J. J., & O'Keefe, J. H. (2017). Good fats versus bad fats: A comparison of fatty acids in the promotion of insulin resistance, inflammation, and obesity. *Missouri medicine, 114*(4), 303-307. Retrieved from https://pubmed.ncbi.nlm.nih.gov/30228616

Dominguez-Salas, P., Galiè, A., Omore, A., Omosa, E., & Ouma, E. (2019). Contributions of milk production to food and nutrition security. In P. Ferranti, E. M. Berry, & J. R. Anderson (Eds.), *Encyclopedia of Food Security and Sustainability* (pp. 278-291). Oxford: Elsevier.

Dona, A. C., Kyriakides, M., Scott, F., Shephard, E. A., Varshavi, D., Veselkov, K., & Everett, J. R. (2016). A guide to the identification of metabolites in NMR-based metabonomics/metabolomics experiments. *Computational and Structural Biotechnology Journal, 14*, 135-153. doi:https://doi.org/10.1016/j.csbj.2016.02.005

Doole, G. J. (2014). Economic feasibility of supplementary feeding on dairy farms in the Waikato region of New Zealand. *New Zealand Journal of Agricultural Research, 57*(2), 90-99. doi:10.1080/00288233.2013.870915

Dos Santos, A. C. S., Santos, S. A., Carvalho, G. G. P., Mariz, L. D. S., Tosto, M. S. L., Valadares Filho, S. C., & Azevedo, J. A. G. (2018). A comparative study on the excretion of urinary metabolites in goats and sheep to evaluate spot sampling applied to protein nutrition trials. *Journal of Animal Science, 96*(8), 3381-3397. doi:10.1093/jas/sky198

Dunn, W. B., Broadhurst, D. I., Atherton, H. J., Goodacre, R., & Griffin, J. L. (2011). Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chemical Society Reviews, 40*(1), 387-426.

Dursun, A., Güler, Z., & Şekerli, Y. E. (2017). Characterization of volatile compounds and organic acids in ultra-high-temperature milk packaged in tetra brik cartons. *International Journal of Food Properties, 20*(7), 1511-1521. doi:10.1080/10942912.2016.1213280

Ellis, D. I., Brewster, V. L., Dunn, W. B., Allwood, J. W., Golovanov, A. P., & Goodacre, R. (2012). Fingerprinting food: current technologies for the detection of food adulteration and contamination. *Chemical Society Reviews, 41*(17), 5706-5727.

Emwas, A.-H., Roy, R., McKay, R. T., Tenori, L., Saccenti, E., Gowda, G. A. N., . . . Wishart, D. S. (2019). NMR Spectroscopy for Metabolomics Research. *Metabolites, 9*(7), 123. doi:10.3390/metabo9070123

Esbensen, K., & Geladi, P. (2009). Principal component analysis: Concept, geometrical interpretation, mathematical background, algoritms, history, practice. In *Comprehensive Chemometrics* (Vol. 2, pp. 211-226).

Everstine, K., Spink, J., & Kennedy, S. (2013). Economically motivated adulteration (EMA) of food: Common characteristics of EMA incidents. *Journal of Food Protection, 76*(4), 723-735.

Fan, T. W.-M., & Lane, A. N. (2016). Applications of NMR spectroscopy to systems biochemistry. *Progress in nuclear magnetic resonance spectroscopy, 92*, 18-53.

FAO. (2020a). *Dairy Market Review*. Retrieved from http://www.fao.org/3/ca8341en/CA8341EN.pdf

FAO. (2020b). *Policy Support and Governance Gatewat: Sustainable Livestock*. Retrieved from http://www.fao.org/policy-support/policy-themes/sustainable-livestock/en/

FAOSTAT. (2014). *Production Livestock Processed Commodities All Data (Milk - All Countries)*. Retrieved from: http://fenixservices.fao.org/faostat/static/bulkdownloads/Production_LivestockProcessed_E_All_Data.zip

FAOSTAT. (2018a). *Number of livestock, goats (World, Total)*.

FAOSTAT. (2018b). *Number of Milk production ,whole fresh goats*. Retrieved from: http://www.fao.org/faostat/en/#data/QA

FAOSTAT. (2018c). *Production livestock primary, sheep milk whole*. Retrieved from: http://www.fao.org/faostat/en/?#data/QL

Farrell, H. M., Jr., Jimenez-Flores, R., Bleck, G. T., Brown, E. M., Butler, J. E., Creamer, L. K., Swaisgood, H. E. (2004). Nomenclature of the proteins of cows' milk--sixth revision. *Journal of Dairy Science, 87*(6), 1641-1674. doi:10.3168/jds.S0022-0302(04)73319-6

Faye, B., & Konuspayeva, G. (2012). The sustainability challenge to the dairy sector – The growing importance of non-cattle milk production worldwide. *International Dairy Journal, 24*, 50-56. doi:10.1016/j.idairyj.2011.12.011

Feligini, M., Bonizzi, I., Cubric Curik, V., Parma, P., Greppi, G. F., & Enne, G. (2005). Detection of adulteration in Italian mozzarella cheese using mitochondrial DNA templates as biomarkers. *Food Technology and Biotechnology, 43*, 91-95.

Fenyvessy, J., Sirokman, K., & Varro, G. (1991). Fatty acid and amino acid composition of sheep milk. *Elelmezesi Ipar (Hungary)*.

Finehout, E. J., & Lee, K. H. (2004). An introduction to mass spectrometry applications in biological research. *Biochemistry and Molecular Biology Education, 32*(2), 93-100. doi:10.1002/bmb.2004.494032020331

Finete, V. d. L. M., Gouvêa, M. M., de Carvalho Marques, F. F., & Netto, A. D. P. (2013). Is it possible to screen for milk or whey protein adulteration with melamine, urea and ammonium sulphate, combining Kjeldahl and classical spectrophotometric methods? *Food Chemistry, 141*(4), 3649-3655.

Finoulst, I., Pinkse, M., Van Dongen, W., & Verhaert, P. (2011). Sample preparation techniques for the untargeted LC-MS-based discovery of peptides in complex biological matrices. *Journal of Biomedicine & Biotechnology, 2011*, 245291-245291. doi:10.1155/2011/245291

Fischer, W. J., Schilter, B., Tritscher, A., & Stadler, R. (2015). Contaminants of milk and dairy products: Contamination resulting from farm and dairy practices. In *Encyclopedia of Dairy Science (Second Edition)* (pp. 1-13).

Flachowsky, G., Meyer, U., & Südekum, K.-H. (2017). Land use for edible protein of animal origin-A review. *Animals, 7*(3), 25. doi:10.3390/ani7030025

Floris, R., Lambers, T., Alting, A., & Kiers, J. (2010). 17 - Trends in infant formulas: a dairy perspective. In M. W. Griffiths (Ed.), *Improving the Safety and Quality of Milk* (pp. 454-474).Cambridge: Woodhead Publishing.

Foegeding, E. A., Davis, J. P., Doucet, D., & McGuffey, M. K. (2002). Advances in modifying and understanding whey protein functionality. *Trends in Food Science & Technology, 13*(5), 151-159. doi:https://doi.org/10.1016/S0924-2244(02)00111-5

Fu, J., & Nicoll, A. (2016). *The Milk Scandal and Corporate Governance in China*.

Fuhrer, T., & Zamboni, N. (2015). High-throughput discovery metabolomics. *Current opinion in biotechnology, 31*, 73-78.

Gabriels, G., Lambert, M., Smith, P., Wiesner, L., & Hiss, D. (2015). Melamine contamination in nutritional supplements--Is it an alarm bell for the general consumer, athletes, and 'Weekend Warriors'? *Nutrition Journal, 14*, 69-69. doi:10.1186/s12937-015-0055-7

Gambelli, L. (2017). Milk and Its Sugar-Lactose: A Picture of Evaluation Methodologies. *Beverages, 3*, 35. doi:10.3390/beverages3030035

Gantner, V., Mijic, P., Baban, M., Škrtić, Z., & Turalija, A. (2015). The overall and fat composition of milk of various species. *Mljekarstvo / Dairy, 65*, 223-231. doi:10.15567/mljekarstvo.2015.0401

García-Cañas, V., & Cifuentes, A. (2008). Recent advances in the application of capillary electromigration methods for food analysis. *Electrophoresis, 29*(1), 294-309.

García-Cañas, V., Simo, C., Castro-Puyana, M., & Cifuentes, A. (2014). Recent advances in the application of capillary electromigration methods for food analysis and Foodomics. *Electrophoresis, 35*(1), 147-169.

Garcia, C., Lutz, N. W., Confort-Gouny, S., Cozzone, P. J., Armand, M., & Bernard, M. (2012). Phospholipid fingerprints of milk from different mammalians determined by 31P NMR: Towards specific interest in human health. *Food Chemistry, 135*(3), 1777-1783. doi:https://doi.org/10.1016/j.foodchem.2012.05.111

Garrote, G. L., Abraham, A. G., & Rumbo, M. (2015). Is lactate an undervalued functional component of fermented food products? *Frontiers in Microbiology, 6*, 629-629. doi:10.3389/fmicb.2015.00629

Gertsman, I., & Barshop, B. A. (2018). Promises and pitfalls of untargeted metabolomics. *Journal of Inherited Metabolic Disease, 41*(3), 355-366. doi:10.1007/s10545-017-0130-7

Getaneh, G., Mebrat, A., Wubie, A., & Kendie, H. (2016). Review on goat milk composition and its nutritive value. *Journal of Nutrition and Health Sciences, 3*(4), 1-10.

Gijsbers, L., Ding, E. L., Malik, V. S., De Goede, J., Geleijnse, J. M., & Soedamah-Muthu, S. S. (2016). Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies.*The American Journal of Clinical Nutrition, 103*(4), 1111-1124.

Gil, A., & Sanchez-Medina, F. (1981). Acid-soluble nucleotides of cow's, goat's and sheep's milks, at different stages of lactation. *Journal of Dairy Research, 48*(1), 35-44. doi:10.1017/S0022029900021427

Godoy, J. L., Vega, J. R., & Marchetti, J. L. (2014). Relationships between PCA and PLS-regression. *Chemometrics and Intelligent Laboratory Systems, 130*, 182-191. doi:https://doi.org/10.1016/j.chemolab.2013.11.008

Goldrick, S., Sandner, V., Cheeks, M., Turner, R., Farid, S. S., McCreath, G., & Glassey, J. (2020). Multivariate data analysis methodology to solve data challenges related to scale-up model validation and missing data on a micro-bioreactor system. *Biotechnology Journal, 15*(3), 1800684. doi:10.1002/biot.201800684

Gómez-Cortés, P., Juárez, M., & de la Fuente, M. A. (2018). Milk fatty acids and potential health benefits: An updated vision. *Trends in Food Science & Technology, 81*, 1-9. doi:https://doi.org/10.1016/j.tifs.2018.08.014

Górska-Warsewicz, H., Rejman, K., Laskowski, W., & Czeczotko, M. (2019). Milk and dairy products and their nutritional contribution to the average Polish diet. *Nutrients, 11*(8), 1771. doi:10.3390/nu11081771

Gossner, C. M.-E., Schlundt, J., Ben Embarek, P., Hird, S., Lo-Fo-Wong, D., Beltran, J. J. O., Tritscher, A. (2009). The melamine incident: implications for international food and feed safety. *Environmental Health Perspectives, 117*(12), 1803-1808. doi:10.1289/ehp.0900949

Gowen, A. A., Downey, G., Esquerre, C., & O'Donnell, C. P. (2011). Preventing over-fitting in PLS calibration models of near-infrared (NIR) spectroscopy data using regression coefficients. *Journal of Chemometrics, 25*(7), 375-381. doi:10.1002/cem.1349

Graulet, B. (2014). Ruminant milk: A source of vitamins in human nutrition. *Animal Frontiers, 4*(2), 24-30.

Graulet, B., & Girard, C. L. (2017). Chapter 15 - B Vitamins in cow milk: Their relevance to human health. In R. R. Watson, R. J. Collier, & V. R. Preedy (Eds.), *Dairy in Human Health and Disease Across the Lifespan* (pp. 211-224): Academic Press.

Graulet, B., Martin, B., Agabriel, C., & Girard, C. L. (2013). Vitamins in milks. *Milk and Dairy Products in Human Nutrition*, 200-219. Retrieved from https://doi.org/10.1002/9781118534168.ch10

Grauwet, T., Vervoort, L., Colle, I., Van Loey, A., & Hendrickx, M. (2014). From fingerprinting to kinetics in evaluating food quality changes. *Trends in Biotechnology, 32*(3), 125-131. doi:https://doi.org/10.1016/j.tibtech.2014.01.002

Griffith, L. (2015). *Business plan for the NZ sheep dairy industry*. Retrieved from https://www.nzagrifoodweek.co.nz/wp-content/uploads/2015/10/2014_BUSINESS-PLAN-FOR-THE-NZ-SHEEP-DAIRY-INDUSTRY_Lucy_Griffiths.pdf

Griffiths, P., & Haseth, J. (2006). Fourier transform infrared spectrometry, Second Edition. doi:10.1002/9780470106310.ch9

Guebel, D. V., & Torres, N. V. (2013). Partial least-squares regression (PLSR). In W. Dubitzky, O. Wolkenhauer, K.-H. Cho, & H. Yokota (Eds.), *Encyclopedia of Systems Biology* (pp. 1646-1648). New York, NY: Springer New York.

Guennec, A. L., Giraudeau, P., & Caldarelli, S. (2014). Evaluation of fast 2D NMR for metabolomics. *Analytical Chemistry, 86*(12), 5946-5954.

Gurevich, V., Kotharu, P., McCann, K., & Bertolini, J. (2017). Improvement of ELISA procedures through simultaneous addition of antigen and detection antibody and elimination of washing steps. *Journal of Immunoassay and Immunochemistry, 38*(5), 494-504. doi:10.1080/15321819.2017.1331171

Haas, R., Schnepps, A., Pichler, A., & Meixner, O. (2019). Cow milk versus plant-based milk substitutes: A comparison of product image and motivational structure of consumption. *Sustainability, 11*(18), 5046.

Haenlein, G. F. (2001). Past, present, and future perspectives of small ruminant dairy research. *Journal of Dairy Science, 84*(9), 2097-2115. doi:10.3168/jds.S0022-0302(01)74655-3

Haenlein, G. F. (2004). Goat milk in human nutrition. *Small Ruminant Research, 51*, 155-163. Retrieved from http://www.capridairyworld.org/files/pdf/Haenlein.goatmilk.human.nutr.pdf

Hall, N. (2019). Why sheep might be the new cows: Everything to know about the growing sheep milk industry. Retrieved from https://thisnzlife.co.nz/why-sheep-might-be-the-new-cows-everything-to-know-about-the-growing-sheep-milk-industry/

Handford, C. E., Campbell, K., & Elliott, C. T. (2016). Impacts of milk fraud on food safety and nutrition with special emphasis on developing countries. *Comprehensive Reviews in Food Science and Food Safety, 15*(1), 130-142.

Harry, L. J. (1977). Comparative observations on the ontogeny of taste preference. *Taste and development: The genesis of sweet preference*(32), 99.

Hazra, T., Sharma, V., Sharma, R., & Arora, S. (2017). A species specific simplex polymerase chain reaction-based approach for detection of goat tallow in heat clarified milk fat (ghee). *International Journal of Food Properties, 20*(sup1), S69-S75. doi:10.1080/10942912.2017.1289542

He, Y., Gao, B., Sophian, A., & Yang, R. (2017).Coil-based rectangular PEC sensors for defect classification. In Y. He, B. Gao, A. Sophian, & R. Yang (Eds.), *Transient Electromagnetic-Thermal Nondestructive Testing* (pp. 55-90): Butterworth-Heinemann.

Héberger, K. (2008).Chemoinformatics—multivariate mathematical–statistical methods for data evaluation. In K. Vékey, A. Telekes, & A. Vertes (Eds.), *Medical Applications of Mass Spectrometry* (pp. 141-169). Amsterdam: Elsevier.

Hochwallner, H., Schulmeister, U., Swoboda, I., Spitzauer, S., & Valenta, R. (2014). Cow's milk allergy: from allergens to new forms of diagnosis, therapy and prevention. *Methods (San Diego, Calif.), 66*(1), 22-33. doi:10.1016/j.ymeth.2013.08.005

Holmes-McNary, M. Q., Cheng, W.-L., Mar, M.-H., Fussell, S., & Zeisel, S. H. (1996). Choline and choline esters in human and rat milk and in infant formulas. *The American Journal of Clinical Nutrition, 64*(4), 572-576.

Hoppel, C. (2003). The role of carnitine in normal and altered fatty acid metabolism. *American Journal of Kidney Diseases, 41*, S4-S12. doi:https://doi.org/10.1016/S0272-6386(03)00112-4

Hornickova, S., Dragounová, H., Hejtmánková, K., Michlova, T., & Hejtmánková, A. (2015). Production of benzoic acid in fermented goat's and sheep's nilk. *Scientia Agriculturae Bohemica, 45*. doi:10.1515/sab-2015-0005

House, J. D., Neufeld, J., & Leson, G. (2010). Evaluating the quality of protein from hemp seed (Cannabis sativa L.) products through the use of the protein digestibility-corrected amino

acid score method. *Journal of Agricultural and Food Chemistry, 58*(22), 11801-11807. doi:10.1021/jf102636b

Hsieh, C.-C., Hernández-Ledesma, B., Fernández-Tomé, S., Weinborn, V., Barile, D., & de Moura Bell, J. M. L. N. (2015). Milk Proteins, peptides, and oligosaccharides: Effects against the 21st Century Disorders. *BioMed Research International, 2015*, 146840. doi:10.1155/2015/146840

Hu, F., Furihata, K., Ito-Ishida, M., Kaminogawa, S., & Tanokura, M. (2004). Nondestructive observation of bovine milk by NMR spectroscopy: analysis of existing States of compounds and detection of new compounds. *Journal of Agricultural and Food Chemistry, 52*(16), 4969-4974. doi:10.1021/jf049616o

Hu, F., Furihata, K., Kato, Y., & Tanokura, M. (2007). Nondestructive Quantification of Organic Compounds in Whole Milk without Pretreatment by Two-Dimensional NMR Spectroscopy. *Journal of Agricultural and Food Chemistry, 55*(11), 4307-4311. doi:10.1021/jf062803x

Huang, S., Yang, D., Yongxin, G., & Zhang, X. (2015). Combined supervised information with PCA via discriminative component selection. *Information Processing Letters, 115*(11), 812-816. doi:https://doi.org/10.1016/j.ipl.2015.06.010

Huang, Y. F., Huang, C. C., Hu, C. C., & Chang, H. T. (2006). Capillary electrophoresis-based separation techniques for the analysis of proteins. *Electrophoresis, 27*(18), 3503-3522.

Hurley, I. P., Coleman, R. C., Ireland, H. E., & Williams, J. H. H. (2004). Measurement of bovine IgG by indirect competitive ELISA as a means of detecting milk adulteration. *Journal of Dairy Science, 87*(3), 543-549. doi:https://doi.org/10.3168/jds.S0022-0302(04)73195-1

Idamokoro, E. M., Gunya, B., & Aliber, M. (2019). Farmers' perception and willingness to consume goat milk and goat milk products: A case study of the central Eastern Cape, South Africa. *Pastoralism, 9*(1), 3. doi:10.1186/s13570-019-0139-7

Iqbal, F. (2017). Chapter 17 - Milk Adulteration: A growing health hazard in Pakistan. In R. R. Watson, R. J. Collier, & V. R. Preedy (Eds.), *Nutrients in Dairy and their Implications on Health and Disease* (pp. 215-222): Academic Press.

Jandal, J. M. (1996). Comparative aspects of goat and sheep milk. *Small Ruminant Research, 22*(2), 177-185. doi:https://doi.org/10.1016/S0921-4488(96)00880-2

Jenkins, T. C., & McGuire, M. A. (2006). Major advances in nutrition: Impact on milk composition. *Journal of Dairy Science, 89*(4), 1302-1310. doi:https://doi.org/10.3168/jds.S0022-0302(06)72198-1

Jenness, R. (1980). Composition and characteristics of goat milk: Review 1968−19791. *Journal of Dairy Science, 63*(10), 1605-1630. doi:https://doi.org/10.3168/jds.S0022-0302(80)83125-0

Johnson, J. M., & Conforti, F. D. (2003). LACTOSE. In B. Caballero (Ed.), *Encyclopedia of Food Sciences and Nutrition* (Second Edition) (pp. 3472-3476). Oxford: Academic Press.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374*(2065), 20150202.

Jonsson, P., Bruce, S. J., Moritz, T., Trygg, J., Sjöström, M., Plumb, R., . . . Holmes, E. (2005). Extraction, interpretation and validation of information for comparing samples in metabolic LC/MS data sets. *Analyst, 130*(5), 701-707.

Jung, Y.-K., Jhon, D.-Y., Kim, K., & Hong, Y.-H. (2011). Quantitative detection of cow milk in goat milk mixtures by real-time PCR. *Korean Journal for Food Science of Animal Resources, 31*, 827-833. doi:10.5851/kosfa.2011.31.6.827

Kalkwarf, H. J., Khoury, J. C., & Lanphear, B. P. (2003). Milk intake during childhood and adolescence, adult bone density, and osteoporotic fractures in US women. *Am J Clin Nutr, 77*(1), 257-265.

Kalyankar, S. D., Khedkar, C. D., Patil, A. M., & Deosarkar, S. S. (2016). Milk: Sources and Composition. In B. Caballero, P. M. Finglas, & F. Toldrá (Eds.), *Encyclopedia of Food and Health* (pp. 741-747). Oxford: Academic Press.

Kalyankar, S. D., Sarode, A., Khedkar, C. D., Deosarkar, S. S., & Rd, P. (2016). Sheep milk. In *Encyclopedia of Food and Health* (pp. 758-763). Oxford: Elsevier.

Kamal, M., & Karoui, R. (2015). Analytical methods coupled with chemometric tools for determining the authenticity and detecting the adulteration of dairy products: A review. *Trends in Food Science & Technology, 46*(1), 27-48. doi:https://doi.org/10.1016/j.tifs.2015.07.007

Kamthania, M., Saxena, J., Saxena, K., & Sharma, D. (2014). Milk adulteration: Methods of detection & remedial measures. *International Journal of Engineering and Technical Research, 1*, 15-20.

Katajamaa, M., & Oresic, M. (2007). Data processing for mass spectrometry-based metabolomics. *Journal of Chromatography A, 1158*(1-2), 318-328. doi:10.1016/j.chroma.2007.04.021

Kawasaki, K., Lafont, A. G., & Sire, J. Y. (2011). The evolution of milk casein genes from tooth genes before the origin of mammals. *Molecular Biology and Evolution, 28*(7), 2053-2061. doi:10.1093/molbev/msr020

Kebede, B., Grauwet, T., Mutsokoti, L., Palmers, S., Vervoort, L., Hendrickx, M., & Van Loey, A. (2014). Comparing the impact of high pressure high temperature and thermal sterilization on the volatile fingerprint of onion, potato, pumpkin and red beet. *Food Research International, 56*, 218-225. doi:https://doi.org/10.1016/j.foodres.2013.12.034

Kebede, B., Ting, V., Eyres, G., & Oey, I. (2020). Volatile changes during storage of shelf stable apple juice: Integrating GC-MS fingerprinting and chemometrics. *Foods, 9*(2), 165.

Keerthi Vasan, K., & Surendiran, B. (2016). Dimensionality reduction using Principal Component Analysis for network intrusion detection. *Perspectives in Science, 8*, 510-512. doi:https://doi.org/10.1016/j.pisc.2016.05.010

Khan, K. M., Krishna, H., Majumder, S. K., & Gupta, P. K. (2015). Detection of urea adulteration in milk using near-infrared Raman spectroscopy. *Food Analytical Methods, 8*(1), 93-102.

Kleckner, I. R., & Foster, M. P. (2011). An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et Biophysica Acta, 1814*(8), 942-968. doi:10.1016/j.bbapap.2010.10.012

Klein, M. S., Almstetter, M. F., Schlamberger, G., Nürnberger, N., Dettmer, K., Oefner, P. J., & Gronwald, W. (2010). Nuclear magnetic resonance and mass spectrometry-based milk metabolomics in dairy cows during early and late lactation. *Journal of Dairy Science, 93*(4), 1539-1550. doi:https://doi.org/10.3168/jds.2009-2563

Klein, M. S., Buttchereit, N., Miemczyk, S. P., Immervoll, A. K., Louis, C., Wiedemann, S., & Gronwald, W. (2012). NMR metabolomic analysis of dairy cows reveals milk glycerophosphocholine to phosphocholine ratio as prognostic biomarker for risk of ketosis. *Journal of proteome research, 11*(2), 1373-1381. doi:10.1021/pr201017n

Knopfler, M. (2016). How compatible is cow's milk with the human immune system? *The Science Journal of the Lander College of Arts and Sciences, 9*(2). Retrieved from https://touroscholar.touro.edu/cgi/viewcontent.cgi?article=1055&context=sjlcas

Koek, M. M., Jellema, R. H., van der Greef, J., Tas, A. C., & Hankemeier, T. (2011). Quantitative metabolomics based on gas chromatography mass spectrometry: status and perspectives. *Metabolomics, 7*(3), 307-328. doi:10.1007/s11306-010-0254-3

Kohler, U., & Luniak, M. (2005). Data inspection using biplots. *The Stata Journal, 5*(2), 208-223.

Kompan, D., & Komprej, A. (2012). The effect of fatty acids in goat milk on health. In *Milk Production-An Up-to-Date Overview of Animal Nutrition, Management and Health*: IntechOpen.

Kopf-Bolanz, K., & Sousa, A. (2017). Nutritional implications of an increasing consumption of bon-dairy plant-based beverages instead of cow's milk in Switzerland. *Journal of Advances in Dairy Research, 05*. doi:10.4172/2329-888X.1000197

Kotu, V., & Deshpande, B. (2019).Introduction. In V. Kotu & B. Deshpande (Eds.), *Data Science (Second Edition)* (pp. 1-18). USA: Morgan Kaufmann.

Krishnan, P., Kruger, N. J., & Ratcliffe, R. G. (2004). Metabolite fingerprinting and profiling in plants using NMR. *Journal of Experimental Botany, 56*(410), 255-265. doi:10.1093/jxb/eri010

Kuhara, T., Ohse, M., Inoue, Y., & Cooper, A. J. (2011). A GC/MS-based metabolomic approach for diagnosing citrin deficiency. *Analytical and Bioanaytical Chemistry, 400*(7), 1881-1894. doi:10.1007/s00216-011-4766-0

Kunz, C., Rudloff, S., Baier, W., Klein, N., & Strobel, S. (2000). Oligosaccharides in human milk: structural, functional, and metabolic aspects. *Annual Review of Nutrition, 20*(1), 699-722.

Lam, H. Y., van Hoffen, E., Michelsen, A., Guikers, K., van der Tas, C. H., Bruijnzeel-Koomen, C. A., & Knulst, A. C. (2008). Cow's milk allergy in adults is rare but severe: both casein and whey proteins are involved. *Clinical and Experimental Allergy, 38*(6), 995-1002. doi:10.1111/j.1365-2222.2008.02968.x

Lamanna, R., Braca, A., Di Paolo, E., & Imparato, G. (2011). Identification of milk mixtures by 1H NMR profiling. *Magnetic Resonance in Chemistry, 49*(S1), S22-S26. doi:https://doi.org/10.1002/mrc.2807

Lämmerhofer, M., & Weckwerth, W. (2013). *Metabolomics in practice: successful strategies to generate and analyze metabolic data*. USA: John Wiley & Sons.

Lange, M., Lee, H., Dallas, D., Le Parc, A., de Moura Bell, J. M. L. N., & Barile, D. (2014). Determining functional properties and sources of recently identified bioactive food components: Oligosaccharides, glycolipids, glycoproteins, and peptides. In N. K. Van Alfen (Ed.), *Encyclopedia of Agriculture and Food Systems* (pp. 441-461). Oxford: Academic Press.

Lankatillake, C., Huynh, T., & Dias, D. (2019). Understanding glycaemic control and current approaches for screening antidiabetic natural products from evidence-based medicinal plants. *Plant Methods*. doi:10.1186/s13007-019-0487-8

Lankin, V. S. (1997). Domestication-related behavior in sheep. *Genetika, 33*(8), 1119-1125.

Lawless, H. T., & Heymann, H. (2013). *Sensory evaluation of food: principles and practices*: Springer Science & Business Media.

Ledgard, S., Judge, A., Smeaton, D., & Boyes, M. (2010). Greenhouse gas emissions from Rotorua dairy farms: Summary report (Report to Ministry of Agriculture and Forestry). *Hamilton, New Zealand: AgResearch*.

Lee, L. C., Liong, C.-Y., & Jemain, A. A. (2018). Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst, 143*(15), 3526-3539. doi:10.1039/C8AN00599K

Leung, A. K., & Sauve, R. S. (2003). Whole cow's milk in infancy. *Paediatrics & Child Health, 8*(7), 419-421. doi:10.1093/pch/8.7.419

Lever, J., Krzywinski, M., & Altman, N. (2017). Principal component analysis. *Nature Methods, 14*(7), 641-642. doi:10.1038/nmeth.4346

Li, Q., Yu, Z., Zhu, D., Meng, X., Pang, X., Liu, Y., & Chen, G. (2017). The application of NMR-based milk metabolite analysis in milk authenticity identification. *Journal of the Science of Food and Agriculture, 97*(9), 2875-2882. doi:10.1002/jsfa.8118

Li, Q., Zhao, Y., Zhu, D., Pang, X., Liu, Y., Frew, R., & Chen, G. (2017). Lipidomics profiling of goat milk, soymilk and bovine milk by UPLC-Q-Exactive Orbitrap Mass Spectrometry. *Food Chemistry, 224*, 302-309. doi:https://doi.org/10.1016/j.foodchem.2016.12.083

Li, Y., Ruan, Q., Li, Y., Ye, G., Lu, X., Lin, X., & Xu, G. (2012). A novel approach to transforming a non-targeted metabolic profiling method to a pseudo-targeted method using the retention time locking gas chromatography/mass spectrometry-selected ions monitoring.

Journal of Chromatography A, 1255, 228-236. doi:https://doi.org/10.1016/j.chroma.2012.01.076

Liao, J., Liu, Y. F., Ku, T., Liu, M. H., & Huang, Y. (2017). Qualitative and quantitative identification of adulteration of milk powder using DNA extracted with a novel method. *Journal of Dairy Science, 100*(3), 1657-1663. doi:10.3168/jds.2016-11900

Lichtenstein, A. H. (2013). Fats and Oils. In B. Caballero (Ed.), *Encyclopedia of Human Nutrition (Third Edition)* (pp. 201-208). Waltham: Academic Press.

Lifschitz, C., & Szajewska, H. (2015). Cow's milk allergy: evidence-based diagnosis and management for the practitioner. *European Journal of Pediatrics, 174*(2), 141-150. doi:10.1007/s00431-014-2422-3

Lindon, J. C. (2003). HPLC-NMR-MS: past, present and future. *Drug Discovery Today, 8*(22), 1021-1022. doi:10.1016/s1359-6446(03)02843-5

Liu, N., Parra, H. A., Pustjens, A., Hettinga, K., Mongondry, P., & van Ruth, S. M. (2018). Evaluation of portable near-infrared spectroscopy for organic milk authentication. *Talanta, 184*, 128-135.

Loeschen, D. (2019). The top 10 dairy producing countries in the world. Retrieved from https://www.mixerdirect.com/blogs/mixer-direct-blog/the-top-10-dairy-producing-countries-in-the-world

Loften, J. R., Linn, J. G., Drackley, J. K., Jenkins, T. C., Soderholm, C. G., & Kertz, A. F. (2014). Invited review: Palmitic and stearic acid metabolism in lactating dairy cows. *Journal of Dairy Science, 97*(8), 4661-4674. doi:https://doi.org/10.3168/jds.2014-7919

Lomer, M., Parkes, G., & Sanderson, J. (2008). lactose intolerance in clinical practice–myths and realities. *Alimentary Pharmacology & Therapeutics, 27*(2), 93-103.

López-Calleja, I., González, I., Fajardo, V., Martín, I., Hernández, P. E., García, T., & Martín, R. (2005). Application of Polymerase Chain Reaction to detect adulteration of sheep's milk with goats' milk. *Journal of Dairy Science, 88*(9), 3115-3120. doi:https://doi.org/10.3168/jds.S0022-0302(05)72993-3

Lopez-Lozano, S., Scholtens, M., Smith, R., Burt, D., Harper, L., Thomas, D., & Tozer, P. (2017). Assessing the current value of milk, meat and fibre products from the goat industry in New Zealand. NZ: New Zealand Society of Animal Production

Lopez, C., Cauty, C., & Guyomarc'h, F. (2019). Unraveling the complexity of milk fat globules to tailor bioinspired emulsions providing health benefits: The key role played by the biological membrane. *European Journal of Lipid Science and Technology, 121*(1), 1800201. doi:10.1002/ejlt.201800201

Lordan, R., Tsoupras, A., Mitra, B., & Zabetakis, I. (2018). Dairy fats and cardiovascular disease: Do we really need to be concerned? *Foods, 7*(29). Retrieved from https://www.mdpi.com/2304-8158/7/3/29/htm

Lovreglio, R., & Ouahiba, S. M. (2014). Goat grazing as a wildfire prevention tool: A basic review. *iForest - Biogeosciences and Forestry, 7*. doi:10.3832/ifor1112-007

Lucey, J. A., & Horne, D. S. (2009). Milk Salts: Technological Significance. In P. McSweeney & P. F. Fox (Eds.), *Advanced Dairy Chemistry: Volume 3: Lactose, Water, Salts and Minor Constituents* (pp. 351-389). New York, NY: Springer New York.

Lutter, P., Savoy-Perroud, M.-C., Campos-Gimenez, E., Meyer, L., Goldmann, T., Bertholet, M.-C., & Delatour, T. (2011). Screening and confirmatory methods for the determination of melamine in cow's milk and milk-based powdered infant formula: Validation and proficiency-tests of ELISA, HPLC-UV, GC-MS and LC-MS/MS. *Food Control, 22*(6), 903-913. doi:https://doi.org/10.1016/j.foodcont.2010.11.022

Mabood, F., Hussain, J., Moo, A., Gilani, S., Farooq, S., Naureen, Z., . . . Al-Harrasi, A. (2016). Detection and quantification of formalin adulteration in cow milk using near infrared spectroscopy combined with multivariate analysis. *Advances in Dairy Research, 05*. doi:10.4172/2329-888X.1000167

Mack, P. B. (1952). *A Preliminary Nutrition Study of the Value of Goat's Milk in the Diet of Children*. Mena, Arkansas, USA: American Goat Society.

Maher, A. D., & Rochfort, S. J. (2014). Applications of NMR in dairy research. *Metabolites, 4*(1), 131-141. doi:10.3390/metabo4010131

Maitra, S., & Yan, J. (2008). Principle component analysis and partial least squares: Two dimension reduction techniques for regression. *Applying Multivariate Statistical Models, 79*, 79-90.

Månsson, H. L. (2008). Fatty acids in bovine milk fat. *Food & Nutrition Research, 52*, 10.3402/fnr.v3452i3400.1821. doi:10.3402/fnr.v52i0.1821

Markiewicz-Keszycka, M., Czyżak-Runowska, G., Lipińska, P., & Wójtowski, J. (2013). Fatty acid profile of milk - A review. *Bulletin- Veterinary Institute in Pulawy, 57*, 135. doi:10.2478/bvip-2013-0026

Martens, J., Berden, G., Bentlage, H., Coene, K. L. M., Engelke, U. F., Wishart, D., & Oomens, J. (2018). Unraveling the unknown areas of the human metabolome: the role of infrared ion spectroscopy. *Journal of Inherited Metabolic Disease, 41*(3), 367-377. doi:10.1007/s10545-018-0161-8

Masoodi, T. A., & Shafi, G. (2010). Analysis of casein alpha S1 & S2 proteins from different mammalian species. *Bioinformation, 4*(9), 430-435. doi:10.6026/97320630004430

Mattar, R., de Campos Mazo, D. F., & Carrilho, F. J. (2012). Lactose intolerance: diagnosis, genetic, and clinical factors. *Clinical and xeperimental Gastroenterology, 5*, 113-121. doi:10.2147/CEG.S32368

McCain, H. R., Kaliappan, S., & Drake, M. A. (2018). Invited review: Sugar reduction in dairy products. *Journal of Dairy Science, 101*(10), 8619-8640. doi:https://doi.org/10.3168/jds.2017-14347

McClements, D. J., Newman, E., & McClements, I. F. (2019). Plant-based milks: A review of the science underpinning their design, fabrication, and performance. *Comprehensive Reviews in Food Science and Food Safety, 18*(6), 2047-2067. doi:https://doi.org/10.1111/1541-4337.12505

McDermott, A., Visentin, G., De Marchi, M., Berry, D. P., Fenelon, M. A., O'Connor, P. M., & McParland, S. (2016). Prediction of individual milk proteins including free amino acids in bovine milk using mid-infrared spectroscopy and their correlations with milk processing characteristics. *Journal of Dairy Science, 99*(4), 3171-3182. doi:https://doi.org/10.3168/jds.2015-9747

McGiven, A. (2016). The future opportunities and challenges for one of the world's largest dairy export firms: Fonterra in New Zealand. *Journal of Applied Business and Economics, 18*(3).

Melfsen, A., Haeussermann, A., & Hartung, E. (2011). Analysis of milk contents based on NIR spectroscopy. *ART-Schriftenreihe*(15), 27-31.

Melzer, N., Wittenburg, D., Hartwig, S., Jakubowski, S., Kesting, U., Willmitzer, L., & Repsilber, D. (2013). Investigating associations between milk metabolite profiles and milk traits of Holstein cows. *Journal of Dairy Science, 96*(3), 1521-1534. doi:10.3168/jds.2012-5743

Ménard, O., Ahmad, S., Rousseau, F., Briard-Bion, V., Gaucheron, F., & Lopez, C. (2010). Buffalo vs cow milk fat globules: Size distribution, zeta-potential, compositions in total fatty acids and in polar lipids from the milk fat globule membrane. *Food Chemistry, 120*, 544-551. doi:10.1016/j.foodchem.2009.10.053

Merlin Junior, I. A., Sifuentes dos Santos, J., Grecco Costa, L., Grecco Costa, R., Ludovico, A., de Almeida Rego, F. C., & Walter de Santana, E. H. (2015). Sheep milk: physical-chemical characteristics and microbiological quality. *Archivos Latinoamericanos de Nutricion, 65*(3), 193-198.

Mermelstein, N. H. (2019). Applying metabolomics to food research. *Food Technology Magazine, 73(2)*. Retrieved from https://www.ift.org/news-and-publications/food-technology-magazine/issues/2019/february/columns/food-safety-and-quality-metabolomics#:~:text=Metabolomics%2C%20she%20said%2C%20can%20be,with%20these%20food%20products%20alter

Meyers, L. D., Hellwig, J. P., & Otten, J. J. (2006). *Dietary reference intakes: the essential guide to nutrient requirements*: National Academies Press.

MFE. (2001). *Valuing New Zealand's clean green image*. Retrieved from https://www.mfe.govt.nz/sites/default/files/clean-green-aug01-final.pdf

Mijan, M., Lee, Y.-K., & Kwak, H.-S. (2011). Classification, structure, and bioactive functions of oligosaccharides in milk. *Korean Journal for Food Science of Animal Resources, 31*, 631-640. doi:10.5851/kosfa.2011.31.5.631

Milani, F. X., & Wendorff, W. L. (2011). Goat and sheep production in the United States (USA). *Small Ruminant Research, 101*. doi:10.1016/j.smallrumres.2011.09.033

Miller, B. A., & Lu, C. D. (2019). Current status of global dairy goat production: an overview. *Asian-Australasian Journal of Animal Sciences, 32*(8), 1219-1232. doi:10.5713/ajas.19.0253

Mobli, M., & Almond, A. (2007). N-Acetylated amino sugars: the dependence of NMR 3J(HNH2)-couplings on conformation, dynamics and solvent. *Organic & Biomolecular Chemistry, 5*(14), 2243-2251. doi:10.1039/B705761J

Mohammad, M. A., Hadsell, D. L., & Haymond, M. W. (2012). Gene regulation of UDP-galactose synthesis and transport: potential rate-limiting processes in initiation of milk production in humans. *The American Journal of Physiology-Endocrinology and Metabolism, 303*(3), E365-376. doi:10.1152/ajpendo.00175.2012

Mohammed, K. H. (2019). Milk Adulteration. *Options to Maintain a Quality Product*: GRIN Verlag.

Monakhova, Y. B., Godelmann, R., Andlauer, C., Kuballa, T., & Lachenmeier, D. W. (2013). Identification of imitation cheese and imitation ice cream based on Vegetable fat using NMR spectroscopy and chemometrics. *International Journal of Food Science, 2013*, 367841. doi:10.1155/2013/367841

Monakhova, Y. B., Kuballa, T., Leitz, J., Andlauer, C., & Lachenmeier, D. W. (2012). NMR spectroscopy as a screening tool to validate nutrition labeling of milk, lactose-free milk, and milk substitutes based on soy and grains. *Dairy Science & Technology, 92*(2), 109-120.

Moncayo, S., Manzoor, S., Navarro-Villoslada, F., & Caceres, J. O. (2015). Evaluation of supervised chemometric methods for sample classification by Laser Induced Breakdown Spectroscopy. *Chemometrics and Intelligent Laboratory Systems, 146*, 354-364. doi:https://doi.org/10.1016/j.chemolab.2015.06.004

Moore, J. C., Spink, J., & Lipp, M. (2012). Development and application of a database of food ingredient fraud and economically.

Moosavy, M., Kordasht, H. K., Khatibi, S., & Sohrabi, H. (2019). Assessment of the chemical adulteration and hygienic quality of raw cow milk in the northwest of Iran. *Quality Assurance and Safety of Crops & Foods, 11*(5), 491-498.

Mozaffarian, D., Hao, T., Rimm, E. B., Willett, W. C., & Hu, F. B. (2011). Changes in diet and lifestyle and long-term weight gain in women and men. *The New England Journal of Medicine, 364*(25), 2392-2404. doi:10.1056/NEJMoa1014296

MPI. (2013). *National Programme for the Monitoring and Surveillance of Chemical Residues and Contaminants in Milk*. Retrieved from https://www.biosecurity.govt.nz/dmsdocument/20210/direct

Food Act 2014, (2018).

MPI. (2019). *Situation and outlook for primary industries*. Wellington, New Zealand: Ministry for Primary Industries.

Muehlhoff, E., Bennett, A., & McMahon, D. (2013). *Milk and dairy products in human nutrition*: Food and Agriculture Organization of the United Nations (FAO).

Mung, D. (2017). Development of chemical isotope labeling LC-MS for milk metabolomics: Comprehensive and quantitative profiling of the amine/phenol submetabolome. *Analytical*

Chemistry, v. 89(no. 8), pp. 4435-4443-2017 v.4489 no.4438. doi:10.1021/acs.analchem.6b03737

Murgia, A., Scano, P., Cacciabue, R., Dessì, D., & Caboni, P. (2019). GC-MS metabolomics comparison of yoghurts from sheep's and goats' milk. *International Dairy Journal*, *96*, 44-49.

Nascimento, C. F., Santos, P. M., Pereira-Filho, E. R., & Rocha, F. R. P. (2017). Recent advances on determination of milk adulterants. *Food Chemistry, 221*, 1232-1244.

Nascimento, W. W. G., Souza, M. P. F., Valente, A. C. M. M., Anjos, V. C., Furtado, M. A. M., & Bell, M. J. V. (2017). Results from portable and of low cost equipment developed for detection of milk adulterations. *Food Science and Technology, 37*, 38-41. Retrieved from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-20612017000700038&nrm=iso

Neumann, C., Harris, D. M., & Rogers, L. M. (2002). Contribution of animal source foods in improving diet quality and function in children in the developing world. *Nutrition research, 22*(1-2), 193-220.

Nicolaou, N., Xu, Y., & Goodacre, R. (2010). Fourier transform infrared spectroscopy and multivariate analysis for the detection and quantification of different milk species. *Journal of Dairy Science, 93*(12), 5651-5660. doi:https://doi.org/10.3168/jds.2010-3619

Nijdam, J. J., & Langrish, T. A. G. (2006). The effect of surface composition on the functional properties of milk powders. *Journal of Food Engineering, 77*(4), 919-925. doi:https://doi.org/10.1016/j.jfoodeng.2005.08.020

NZIER. (2017). *Dairy trade's economic contribution to New Zealand*. Retrieved from https://www.dcanz.com/UserFiles/DCANZ/File/Dairy%20economic%20contribution%20update%20FINAL%2021%20February%202017.pdf

NZMP. (2018). What makes New Zealand milk so good? Retrieved from https://www.nzmp.com/global/en/news/what-makes-nz-milk-so-good.html

O'Kennedy, B. T. (2011). Caseins. In G. O. Phillips & P. A. Williams (Eds.), *Handbook of Food Proteins* (pp. 13-29): Woodhead Publishing.

Odziomek, K., Rybinska, A., & Puzyn, T. (2017). Unsupervised learning methods and similarity analysis in chemoinformatics. In J. Leszczynski, A. Kaczmarek-Kedziera, T. Puzyn, M. G. Papadopoulos, H. Reis, & M. K. Shukla (Eds.), *Handbook of Computational Chemistry* (pp. 2095-2132). Cham: Springer International Publishing.

OEC. (2017). Where does New Zealand export concentrated milk to? In: *The Observatory of Economic Complexity*.

OECD, & FAO. (2016). *OECD-FAO Agricultural Outlook 2016-2025* (OECD Ed.).

OECD, & FAO. (2019a). Chapter 7. Dairy and dairy products. In *OECD-FAO Agricultural Outlook 2019-2028*: OECD Publishing.

OECD, & FAO. (2019b). *OECD-FAO Agricultural Outlook 2019-2028*: OECD.

Oliveri, P., & Forina, M. (2012). Chapter 2 - Data Analysis and Chemometrics. In Y. Picó (Ed.), *Chemical Analysis of Food: Techniques and Applications* (pp. 25-57). Boston: Academic Press.

Ooms, K. (1996). Identification of potentially causal regressors in PLS models. *Catholic University Leuven, Leuven*.

Opio, C., Gerber, P., Mottet, A., Falcucci, A., Tempio, G., MacLeod, M., . . . Steinfeld, H. (2013). *Greenhouse gas emissions from ruminant supply chains: A global life cycle assessment*. Rome: Food and Agriculture Organization of the United Nations (FAO).

Osthoff, G. (2016). Milk: Milks of Non-dairy Mammals. In *Reference Module in Food Science*: Elsevier.

Oyedele, O. F., & Lubbe, S. (2015). The construction of a partial least-squares biplot. *Journal of Applied Statistics, 42*(11), 2449-2460. doi:10.1080/02664763.2015.1043858

Palacios, O., Badran, J., Drake, M., Reisner, M., & Moskowitz, H. (2009). Consumer acceptance of cow's milk versus soy beverages: Impact of ethnicity, lactose tolerance and sensory preference segmentation. *Journal of Sensory Studies, 24*, 731-748. doi:10.1111/j.1745-459X.2009.00236.x

Pan, L., Yu, J., Mi, Z., Mo, L., Jin, H., Yao, C., & Menghe, B. (2018). A metabolomics approach uncovers differences between traditional and commercial dairy products in Buryatia (Russian Federation). *Molecules, 23*(4), 735. Retrieved from https://www.mdpi.com/1420-3049/23/4/735

Paniagua, J. A., de la Sacristana, A. G., Sánchez, E., Romero, I., Vidal-Puig, A., Berral, F. J., & Pérez-Jiménez, F. (2007). A MUFA-rich diet improves posprandial glucose, lipid and GLP-1 responses in insulin-resistant subjects. *Journal of the American College of Nutrition, 26*(5), 434-444. doi:10.1080/07315724.2007.10719633

Panthi, R. R., Jordan, K. N., Kelly, A. L., & Sheehan, J. J. (2017). Selection and Treatment of Milk for Cheesemaking. In P. L. H. McSweeney, P. F. Fox, P. D. Cotter, & D. W. Everett (Eds.), *Cheese (Fourth Edition)* (pp. 23-50). San Diego: Academic Press.

Paradkar, M. M., Singhal, R. S., & Kulkarni, P. R. (2000). An approach to the detection of synthetic milk in dairy milk: 1. Detection of urea. *International Journal of Dairy Technology, 53*(3), 87-91.

Park, S.-J., Jung, J. H., Kim, M.-S., & Lee, H.-J. (2018). High dairy products intake reduces osteoporosis risk in Korean postmenopausal women: A 4 year follow-up study. *Nutrition research and practice, 12*(5), 436-442. doi:10.4162/nrp.2018.12.5.436

Park, Y. W. (2009). *Bioactive components in milk and dairy products*: John Wiley & Sons.

Park, Y. W., Juárez, M., Ramos, M., & Haenlein, G. F. W. (2007). Physico-chemical characteristics of goat and sheep milk. *Small Ruminant Research, 68*(1), 88-113. doi:https://doi.org/10.1016/j.smallrumres.2006.09.013

Parodi, P. W. (2004). Milk fat in human nutrition. *Australian Journal of Dairy Technology, 59*(1), 3.

Peaker, M., & Linzell, J. L. (1975). Citrate in milk: a harbinger of lactogenesis. *Nature, 253*(5491), 464-464. doi:10.1038/253464a0

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2*(11), 559-572. doi:10.1080/14786440109462720

Penn, D., Dolderer, M., & Schmidt-Sommerfeld, E. (1987). Carnitine concentrations in the milk of different species and infant formulas. *Neonatology, 52*(2), 70-79. doi:10.1159/000242686

Peter, S. C., Dhanjal, J. K., Malik, V., Radhakrishnan, N., Jayakanthan, M., & Sundar, D. (2019). Quantitative Structure-Activity Relationship (QSAR): Modeling approaches to biological applications. In S. Ranganathan, M. Gribskov, K. Nakai, & C. Schönbach (Eds.), *Encyclopedia of Bioinformatics and Computational Biology* (pp. 661-676). Oxford: Academic Press.

Peterson, S. W., & Prichard, C. (2015). *The sheep dairy industry in New Zealand: a review*. Paper presented at the Proceedings of the New Zealand Society of Animal Production, Dunedin.

Piantoni, P., Lock, A. L., & Allen, M. S. (2013). Palmitic acid increased yields of milk and milk fat and nutrient digestibility across production level of lactating cows. *Journal of Dairy Science, 96*(11), 7143-7154. doi:https://doi.org/10.3168/jds.2013-6680

Pietrzak-Fiećko, R., & Kamelska-Sadowska, A. M. (2020). The comparison of nutritional value of human milk with other mammals' milk. *Nutrients, 12*(5), 1404.

Pisano, M. B., Scano, P., Murgia, A., Cosentino, S., & Caboni, P. (2016). Metabolomics and microbiological profile of Italian mozzarella cheese produced with buffalo and cow milk. *Food Chemistry, 192*, 618-624. doi:10.1016/j.foodchem.2015.07.061

Pizzo, J. S., Galuch, M. B., Santos, P. D., Santos, O. O., Visentainer, L., Eberlin, M. N., & Visentainer, J. V. (2018). Assessment of adulteration of cosmetics based on vegetable oils by GC-FID and lipid profile using direct infusion Electrospray Ionization Mass Spectrometry (ESI-MS). *Journal of the Brazilian Chemical Society, 29*(12), 2457-2465.

Plakantara, S., Michaelidou, A. M., Polychroniadou, A., Menexes, G., & Alichanidis, E. (2010). Nucleotides and nucleosides in ovine and caprine milk during lactation. *Journal of Dairy Science, 93*(6), 2330-2337. doi:https://doi.org/10.3168/jds.2009-2836

Pomerantsev, A. L. (2008). Acceptance areas for multivariate classification derived by projection methods. *Journal of Chemometrics, 22*(11-12), 601-609. doi:10.1002/cem.1147

Poonia, A., Jha, A., Sharma, R., Singh, H., Rai, A., & Sharma, N. (2016). Detection of adulteration in milk: A review. *International Journal of Dairy Technology*. doi:10.1111/1471-0307.12274

Potočnik, K., Gantner, V., Kuterovac, K., & Angela, C. (2011). Mare's milk: composition and protein fraction in comparison with different milk species. *Mljekarstvo / Dairy, 61*, 107-113.

Prajapati, P., & Agrawal, Y. K. (2014). Analysis and impurity identification in pharmaceuticals. *Reviews in Analytical Chemistry, 33*(2), 123-133. doi:doi:10.1515/revac-2014-0001

Preble, I., Zhang, Z., Kopp, R., Garzotto, M., Bobe, G., Shannon, J., & Takata, Y. (2019). Dairy product consumption and prostate cancer risk in the United States. *Nutrients, 11*(7), 1615. doi:10.3390/nu11071615

Prieto, N., Pawluczyk, O., Dugan, M. E. R., & Aalhus, J. L. (2017). A Review of the Principles and Applications of Near-Infrared Spectroscopy to Characterize Meat, Fat, and Meat Products. *Applied Spectroscopy, 71*(7), 1403-1426. doi:10.1177/0003702817709299

Pulina, G., Milán, M. J., Lavín, M. P., Theodoridis, A., Morin, E., Capote, J., . . . Caja, G. (2018). Invited review: Current production trends, farm structures, and economics of the dairy sheep and goat sectors. *Journal of Dairy Science, 101*(8), 6715-6729. doi:https://doi.org/10.3168/jds.2017-14015

Qian, G., Guo, X., Guo, J., & Wu, J. (2011). China's dairy crisis: impacts, causes and policy implications for a sustainable dairy industry. *International Journal of Sustainable Development & World Ecology, 18*(5), 434-441. doi:10.1080/13504509.2011.581710

Rácz, A., Bajusz, D., & Héberger, K. (2018). Chemometrics in analytical chemistry. In *Applied Chemoinformatics* (pp. 471-499).

Rafiq, S., Huma, N., Pasha, I., Sameen, A., Mukhtar, O., & Khan, M. I. (2016). Chemical Composition, Nitrogen Fractions and Amino Acids Profile of Milk from Different Animal Species. *Asian-Australasian journal of animal sciences, 29*(7), 1022-1028. doi:10.5713/ajas.15.0452

Rako, A., Kalit, M. T., & Kalit, S. (2019). Effect of sheep's milk composition on strength and syneresis of rennet-induced milk gel during lactation. *Food Technology and Biotechnology, 57*(3), 426-433. doi:10.17113/ftb.57.03.19.6218

Ramel, P. R., & Marangoni, A. G. (2019). Microstructure of Dairy Fat Products. In L. Melton, F. Shahidi, & P. Varelis (Eds.), *Encyclopedia of Food Chemistry* (pp. 39-46). Oxford: Academic Press.

Ramos, M., & Juarez, M. (2011). Milk | Sheep Milk. In J. W. Fuquay (Ed.), *Encyclopedia of Dairy Sciences (Second Edition)* (pp. 494-502). San Diego: Academic Press.

Rangel, A. H. d. N., Sales, D. C., Urbano, S. A., Galvão Júnior, J. G. B., Andrade Neto, J. C. d., & Macêdo, C. d. S. (2016). Lactose intolerance and cow's milk protein allergy. *Food Science and Technology, 36*, 179-187. Retrieved from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-20612016000200179&nrm=iso

Rasmussen, N., Andersen, J. H., Jespersen, H., Mouritsen, O. G., & Ditzel, H. J. (2010). Effect of free fatty acids and lysolipids on cellular uptake of doxorubicin in human breast cancer cell lines. *Anticancer Drugs, 21*(7), 674-677. doi:10.1097/CAD.0b013e32833c2cf7

Rassin, D. K., Sturman, J. A., & Guall, G. E. (1978). Taurine and other free amino acids in milk of man and other mammals. *Early Human Development, 2*(1), 1-13. doi:10.1016/0378-3782(78)90048-8

Raynal-Ljutovac, K., Lagriffoul, G., Paccard, P., Guillet, I., & Chilliard, Y. (2008). Composition of goat and sheep milk products: An update. *Small Ruminant Research, 79*(1), 57-72. doi:https://doi.org/10.1016/j.smallrumres.2008.07.009

Razafindrakoto, O., Ravelomanana, N., Rasolofo, A., Rakotoarimanana, R. D., Gourgue, P., Coquin, P., & Desjeux, J.-F. (1994). Goat's milk as a substitute for cow's milk in undernourished children: A randomized double-blind clinical trial. *Pediatrics, 94*(1), 65-69. Retrieved from https://pediatrics.aappublications.org/content/pediatrics/94/1/65.full.pdf

Reddy, D. M., Venkatesh, K., & Reddy, C. V. S. (2017). Adulteration of milk and its detection: A review. *International Journal of Chemical Studies, 5*(4), 613-617.

Reinberg, S. (2008). FDA rules how much melamine is too much. Retrieved from https://abcnews.go.com/Health/Healthday/story?id=5952274&page=1

Ren, Q. R., Zhang, H., Guo, H. Y., Jiang, L., Tian, M., & Ren, F. Z. (2014). Detection of cow milk adulteration in yak milk by ELISA. *Journal of Dairy Science, 97*(10), 6000-6006. doi:10.3168/jds.2014-8127

Ritchie, H., & Roser, M. (2017). Meat and dairy production. *Our World in Data*.

Roberts, L. D., Souza, A. L., Gerszten, R. E., & Clish, C. B. (2012). Targeted metabolomics. *Current protocols in molecular biology, Chapter 30*, Unit30.32-30.32.24. doi:10.1002/0471142727.mb3002s98

Rodríguez-Pérez, R., Fernández, L., & Marco, S. (2018). Overoptimism in cross-validation when using partial least squares-discriminant analysis for omics data: A systematic study. *Analytical and Bioanalytical Chemistry, 410*(23), 5981-5992. doi:10.1007/s00216-018-1217-1

Rozenberg, S., Body, J.-J., Bruyère, O., Bergmann, P., Brandi, M. L., Cooper, C., . . . Reginster, J.-Y. (2016). Effects of dairy products consumption on health: Benefits and beliefs—A commentary from the Belgian bone club and the European society for clinical and economic aspects of osteoporosis, osteoarthritis and musculoskeletal diseases. *Calcified Tissue International, 98*(1), 1-17. doi:10.1007/s00223-015-0062-x

Rugoho, I., Liu, Y., & Dewhurst, R. J. (2014). Analysis of major fatty acids in milk produced from high-quality grazed pasture. *New Zealand Journal of Agricultural Research, 57*(3), 165-179. doi:10.1080/00288233.2014.899505

Ruiz-Perez, D., & Narasimhan, G. (2018). So you think you can PLS-DA? *bioRxiv*, 207225.

Ruska, D., & Jonkus, D. (2014). Crude protein and non-protein nitrogen content in dairy cow milk. *Rural Sustainability Research, 32*(1), 36-40.

Sacchi, R., Paduano, A., Caporaso, N., Picariello, G., Romano, R., & Addeo, F. (2018). Assessment of milk fat content in fat blends by 13C NMR spectroscopy analysis of butyrate. *Food Control, 91*, 231-236.

Salama, A. A., Contreras-Jodar, A., Love, S., Mehaba, N., Such, X., & Caja, G. (2020). Milk yield, milk composition, and milk metabolomics of dairy goats intramammary-challenged with lipopolysaccharide under heat stress conditions. *Scientific Reports, 10*(1), 1-12.

Salih, M., & Yang, S. (2017). Common milk adulteration in developing countries case study in China and Sudan: A review. *Advances in Dairy Research, 5*. doi:10.4172/2329-888X.1000192

Salleh, N. A., Selamat, J., Meng, G. Y., Abas, F., Jambari, N. N., & Khatib, A. (2019). Fourier transform infrared spectroscopy and multivariate analysis of milk from different goat breeds. *International Journal of Food Properties, 22*(1), 1673-1683. doi:10.1080/10942912.2019.1668803

Sammut, C., & Webb, G. I. (2010). Leave-One-Out Cross-Validation. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of Machine Learning* (pp. 600-601). Boston, MA: Springer US.

Sanchez, L. J., Zhu, D., Frew, R., & Kebede, B. (2020). Optimization of nuclear magnetic resonance and gas chromatography-mass spectrometry-based fingerprinting methods to characterize goat milk powder. *Journal of Dairy Science*. doi:https://doi.org/10.3168/jds.2020-18467

Sanon, H. O., Kaboré-Zoungrana, C., & Ledin, I. (2007). Behaviour of goats, sheep and cattle and their selection of browse species on natural pasture in a Sahelian area. *Small Ruminant Research, 67*(1), 64-74. doi:https://doi.org/10.1016/j.smallrumres.2005.09.025

Sargaeva, N., Wilson, J., & Stacey, C. (2014). LC/MS study of casein proteins in milk. *PerkinELmer Application Note*.

Sarı, H., Esen, B., Yıldırım, S., Pilten, S., & Aydın, H. (2019). Serum α-Hydroxybutyrate: A candidate marker of insulin resistance is associated with deterioration in anthropometric measurements in individuals with low diabetes risk. *The Journal of Applied Laboratory Medicine, 1*(5), 562-567. doi:10.1373/jalm.2016.021444

Sarker, S. D., & Nahar, L. (2015). Applications of High Performance Liquid Chromatography(HPLC) in the analysis of herbal products. In P. K. Mukherjee (Ed.), *Evidence-Based Validation of Herbal Medicine* (pp. 405-425). Boston: Elsevier.

Scano, P., Cagliani, L. R., & Consonni, R. (2019). 1H NMR characterisation of the lipid fraction and the metabolite profiles of Fossa (pit) cheese. *International Dairy Journal, 90*, 39-44. doi:https://doi.org/10.1016/j.idairyj.2018.10.007

Scano, P., Carta, P., Ibba, I., Manis, C., & Caboni, P. (2020). An untargeted metabolomic comparison of milk composition from sheep kept under different grazing systems. *Dairy, 1*, 4. doi:10.3390/dairy1010004

Scano, P., Cusano, E., Caboni, P., & Consonni, R. (2019). NMR metabolite profiles of dairy: A review. *International Dairy Journal, 90*, 56-67. doi:https://doi.org/10.1016/j.idairyj.2018.11.004

Scano, P., Murgia, A., Pirisi, F. M., & Caboni, P. (2014). A gas chromatography-mass spectrometry-based metabolomic approach for the characterization of goat milk compared with cow milk. *Journal of Dairy Science, 97*(10), 6057-6066.

Scholtens, M., Lopez-Lozano, R., & Smith, R. (2017). New Zealand Goat Industry. Paper presented at the Proceedings of the New Zealand Society of Animal Production.

Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D., & McLean, J. A. (2016). Untargeted metabolomics strategies-challenges and emerging directions. *Journal of the American Society for Mass Spectrometry, 27*(12), 1897-1905. doi:10.1007/s13361-016-1469-y

Schwingshackl, L., & Hoffmann, G. (2012). Monounsaturated fatty acids and risk of cardiovascular disease: synopsis of the evidence available from systematic reviews and meta-analyses. *Nutrients, 4*(12), 1989-2007. doi:10.3390/nu4121989

Scott, A., Bowden, S., & Rowarth, J. (2013). *Critical success factors when going global: New Zealand dairy companies.* Paper presented at the Proceedings of the New Zealand Grassland Association.

Selvaggi, M., Laudadio, V., Dario, C., & Tufarelli, V. (2014). Investigating the genetic polymorphism of sheep milk proteins: a useful tool for dairy production. *Journal of the Science of Food and Agriculture, 94*(15), 3090-3099. doi:10.1002/jsfa.6750

Sethi, S., Tyagi, S. K., & Anurag, R. K. (2016). Plant-based milk alternatives an emerging segment of functional beverages: a review. *Journal of Food Science and Technology, 53*(9), 3408-3423. doi:10.1007/s13197-016-2328-3

Shadbolt, N. M., & Apparao, D. (2016). Factors influencing the dairy trade from New Zealand. *International Food and Agribusiness Management Review Special Issue, 19*(B), 241-255. Retrieved from https://www.ifama.org/resources/Documents/v19ib/14_Shadbolt.pdf

Sharma, R., Rajput, Y. S., & Barui, A. K. (2012). *Detection of Adulterants in Milk: A Laboratory Manual*. India.

Silanikove, N., Leitner, G., & Merin, U. (2016). Influence of animal health, breed, and diet on non-cow milk composition. In E. Tsakalidou & K. Papadimitriou (Eds.), *Non-Bovine Milk and Milk Products* (pp. 61-79). San Diego: Academic Press.

Sinanoglou, V. J., Koutsouli, P., Fotakis, C., Sotiropoulou, G., Cavouras, D., & Bizelis, I. (2015). Assessment of lactation stage and breed effect on sheep milk fatty acid profile and lipid quality indices. *Dairy Science & Technology, 95*(4), 509-531. doi:10.1007/s13594-015-0234-5

Singh, A., Chandra, G., Aggarwal, A., & Kumar, P. (2012). Adulteration detection in milk. *Res News U, 5*, 52-55.

Singh, B., Ganguly, A., & Sunwoo, H. (2016). Current and future diagnostic tests for Ebola virus disease. *Journal of Pharmacy & Pharmaceutical Sciences, 19*, 530. doi:10.18433/J38C9N

Singh, G. M., Micha, R., Khatibzadeh, S., Shi, P., Lim, S., Andrews, K. G., .Chronic Diseases Expert, G. (2015). Global, regional, and national consumption of sugar-sweetened beverages, fruit juices, and milk: A systematic assessment of beverage intake in 187 countries. *PloS one, 10*(8), e0124845. doi:10.1371/journal.pone.0124845

Singh, P., & Gandhi, N. (2015). Milk preservatives and adulterants: processing, regulatory and safety issues. *Food Reviews International, 31*(3), 236-261.

Skeaff, C. M., Williscroft, K., Mann, J., & Chisholm, A. (2004). Replacing cows' with sheep's dairy fat lowers plasma cholesterol concentration in participants consuming dairy fat-rich diets. *European Journal of Clinical Nutrition, 58*(2), 250-257. doi:10.1038/sj.ejcn.1601774

Skibiel, A. L., Downing, L. M., Orr, T. J., & Hood, W. R. (2013). The evolution of the nutrient composition of mammalian milks. *Journal of Animal Ecology, 82*(6), 1254-1264. doi:10.1111/1365-2656.12095

Smith, L. I. (2002). *A tutorial on principal components analysis*.

Sneddon, N. W., Lopez-Villalobos, N., Hickson, R. E., Shalloo, L., Davis, S. R., Garrick, D. J., & Geary, U. (2015). Predicted dairy product yields and deficits of lactose for manufacturing under differing selection and manufacturing scenarios in New Zealand. *New Zealand Journal of Agricultural Research, 58*(4), 432-440. doi:10.1080/00288233.2015.1073160

Solis-ramirez J, L.-v. N. B. H. T. (2011). *Dairy goat production systems in Waikato, New Zealand* (Vol. 71): New Zealand Society of Animal Production.

Souhassou, S., Bassbasi, M., Hirri, A., Kzaiber, F., & Oussama, A. (2018). Detection of camel milk adulteration using Fourier transformed infrared spectroscopy FT-IR coupled with chemometrics methods. *International Food Research Journal, 25*(3), 1213-1218.

Soulat, J., Andueza, D., Graulet, B., Girard, C. L., Labonne, C., Aït-Kaddour, A., & Ferlay, A. (2020). Comparison of the Potential Abilities of Three Spectroscopy Methods: Near-Infrare, Mid-Infrared, and molecular fluorescence, to predict carotenoid, vitamin and fatty acid contents in cow milk. *Foods (Basel, Switzerland), 9*(5), 592. doi:10.3390/foods9050592

Spink, J. (2014). Safety of food and beverages: Risks of food adulteration. In Y. Motarjemi (Ed.), *Encyclopedia of Food Safety* (pp. 413-416). Waltham: Academic Press.

Statista. (2019). Number of sheep livestock in New Zealand from 2010 to 2019. In. Statistic New Zealand.

StatsNZ. (2019). Agricultural production statistics: June 2019 (provisional). In: Statistic New Zealand.

StatsNZ. (2021). New Zealand resident population, year ended December 2020. In: Statistic New Zealand.

Stuart, B. (2005). Infrared spectroscopy. In *Kirk-Othmer Encyclopedia of Chemical Technology*.

Su, L., Ding, X., Cheng, X., Li, P., Huang, Z., Mao, C., & Hao, M. (2020). Study on the potential chemical markers for the discrimination between raw and processed Schisandrae Chinensis Fructus using UPLC-Q-TOF/MS coupled with multivariate statistical analyses. *Journal of Herbal Medicine, 19*, 100311.

Sugrue, I., Tobin, C., Ross, R. P., Stanton, C., & Hill, C. (2019). Foodborne Pathogens and Zoonotic Diseases. In L. A. Nero & A. F. De Carvalho (Eds.), *Raw Milk* (pp. 259-272): Academic Press.

Sun, F., Liu, L., Kuang, H., & Xu, C. (2013). Development of ELISA for melamine detection in milk powder. *Food and Agricultural Immunology, 24*(1), 79-86. doi:10.1080/09540105.2011.641170

Sundekilde, U., Poulsen, N. A., Larsen, L. B., & Bertram, H. (2013). Nuclear magnetic resonance metabonomics reveals strong association between milk metabolites and somatic cell count in bovine milk. *Journal of Dairy Science, 96*(1), 290-299.

Sundekilde, U. K., Frederiksen, P. D., Clausen, M. R., Larsen, L. B., & Bertram, H. C. (2011). Relationship between the metabolite profile and technological properties of bovine milk from two dairy breeds elucidated by NMR-based metabolomics. *Journal of Agricultural and Food Chemistry, 59*(13), 7360-7367. doi:10.1021/jf202057x

Sundekilde, U. K., Gustavsson, F., Poulsen, N. A., Glantz, M., Paulsson, M., Larsen, L. B., & Bertram, H. C. (2014). Association between the bovine milk metabolome and rennet-induced coagulation properties of milk. *Journal of Dairy Science, 97*(10), 6076-6084. doi:https://doi.org/10.3168/jds.2014-8304

Sundekilde, U. K., Larsen, L. B., & Bertram, H. C. (2013). NMR-based milk metabolomics. *Metabolites, 3*(2), 204-222.

Sundekilde, U. K., Poulsen, N. A., Larsen, L. B., & Bertram, H. C. (2013). Nuclear magnetic resonance metabonomics reveals strong association between milk metabolites and somatic cell count in bovine milk. *Journal of Dairy Science, 96*(1), 290-299.

Swartz, M. E. (2005). UPLC™: an introduction and review. *Journal of Liquid Chromatography & Related Technologies, 28*(7-8), 1253-1263.

Taj Khan, I., Nadeem, M., Imran, M., Ullah, R., Ajmal, M., & Jaspal, M. (2019). Antioxidant properties of milk and dairy products: A comprehensive review of the current knowledge. *Lipids in Health and Disease, 18*. doi:10.1186/s12944-019-0969-8

Takis, P. G., Ghini, V., Tenori, L., Turano, P., & Luchinat, C. (2019). Uniqueness of the NMR approach to metabolomics. *TrAC Trends in Analytical Chemistry, 120*, 115300.

Taleuzzaman, M., Ali, S., Gilani, S., Imam, S., & Hafeez, A. (2015). Ultra performance liquid chromatography (UPLC)-a review. *Austin Journal of Analytical and Pharmaceutical Chemistry, 2*(6), 1056.

Tallapragada, P., & Rayavarapu, B. (2019). 5 - Recent trends and developments in milk-based beverages. In A. M. Grumezescu & A. M. Holban (Eds.), *Milk-Based Beverages* (pp. 139-172): Woodhead Publishing.

Tan, D., Zhang, X., Su, M., Jia, M., Zhu, D., Kebede, B., & Chen, G. (2021). Establishing an untargeted-to-MRM liquid chromatography-mass spectrometry method for discriminating reconstituted milk from ultra-high temperature milk. *Food Chem, 337*, 127946. doi:10.1016/j.foodchem.2020.127946

Tang, L., Peng, S., Bi, Y., Shan, P., & Hu, X. (2014). A new method combining LDA and PLS for dimension reduction. *PloS one, 9*(5), e96944. doi:10.1371/journal.pone.0096944

Tang, X., Bai, Y., Duong, A., Smith, M. T., Li, L., & Zhang, L. (2009). Formaldehyde in China: production, consumption, exposure levels, and health effects. *Environment International, 35*(8), 1210-1224. doi:10.1016/j.envint.2009.06.002

TDB. (2018). *New Zealand Dairy Companies Review*. Retrieved from Wellington, New Zealand: https://www.tdb.co.nz/wp-content/uploads/2018/05/TDB-Dairy-Companies-Review-2018-1.pdf

Tenori, L., Santucci, C., Meoni, G., Morrocchi, V., Matteucci, G., & Luchinat, C. (2018). NMR metabolomic fingerprinting distinguishes milk from different farms. *Food Research International, 113*, 131-139. doi:https://doi.org/10.1016/j.foodres.2018.06.066

Thompson, S. Y., Henry, K. M., & Kon, S. K. (1964). Factors affecting the concentration of vitamins in milk: I. Effect of breed, season and geographical location on fat-soluble vitamins. *Journal of Dairy Research, 31*(1), 1-25. doi:10.1017/S0022029900017866

Thorning, T. K., Raben, A., Tholstrup, T., Soedamah-Muthu, S. S., Givens, I., & Astrup, A. (2016). Milk and dairy products: good or bad for human health? An assessment of the totality of scientific evidence. *Food & Nutrition Research, 60*, 32527-32527. doi:10.3402/fnr.v60.32527

Tian, H., Zheng, N., Wang, W., Cheng, J., Li, S., Zhang, Y., & Wang, J. (2016). Integrated metabolomics study of the milk of heat-stressed lactating dairy cows. *Scientific Reports, 6*(1), 24208. doi:10.1038/srep24208

Tiemeyer, W., Stohrer, M., & Giesecke, D. (1984). Metabolites of nucleic acids in bovine milk. *Journal of Dairy Science, 67*(4), 723-728. doi:https://doi.org/10.3168/jds.S0022-0302(84)81361-2

Toso, B., Procida, G., & Stefanon, B. (2002). Determination of volatile compounds in cows' milk using headspace GC-MS. *The Journal of dairy research, 69*(4), 569.

Trimboli, F., Costanzo, N., Lopreiato, V., Ceniti, C., Morittu, V. M., Spina, A., & Britti, D. (2019). Detection of buffalo milk adulteration with cow milk by capillary electrophoresis analysis. *Journal of Dairy Science, 102*(7), 5962-5970. doi:10.3168/jds.2018-16194

Trimboli, F., Morittu, V. M., Cicino, C., Palmieri, C., & Britti, D. (2017). Rapid capillary electrophoresis approach for the quantification of ewe milk adulteration with cow milk. *Journal of Chromatography A, 1519*, 131-136. doi:https://doi.org/10.1016/j.chroma.2017.08.075

Tripaldi, C., Martillotti, F., & Terramoccia, S. (1998). Content of taurine and other free amino acids in milk of goats bred in Italy. *Small Ruminant Research, 30*(2), 127-136. doi:https://doi.org/10.1016/S0921-4488(98)00095-9

Tsai, S.-R., & Hamblin, M. R. (2017). Biological effects and medical applications of infrared radiation. *Journal of Photochemistry and Photobiology B, Biology, 170*, 197-207. doi:10.1016/j.jphotobiol.2017.04.014

Ulfman, L. H., Leusen, J. H. W., Savelkoul, H. F. J., Warner, J. O., & van Neerven, R. J. J. (2018). Effects of bovine immunoglobulins on immune function, allergy, and infection. *Frontiers in Nutrition, 5*(52). doi:10.3389/fnut.2018.00052

Urashima, T., & Taufik, E. (2010). Oligosaccharides in milk: Their benefits and future utilization. *Media Peternakan, 33*. doi:10.5398/medpet.2010.33.3.189

USDA. (2015). *2015-2020 Dietary Guidelines for American*. Washington D.C, U.S.A: United States Department of Agriculture and United States Department of Health and Human Services Retrieved from https://health.gov/sites/default/files/2019-09/2015-2020_Dietary_Guidelines.pdf

Vanga, S. K., & Raghavan, V. (2018). How well do plant based alternatives fare nutritionally compared to cow's milk? *Journal of Food Science and Technology, 55*(1), 10-20. doi:10.1007/s13197-017-2915-y

Vasconcelos, A., Santos, T., Ravasco, P., & Neves, P. M. (2019). Dairy products: Is there an impact on promotion of prostate cancer? A review of the literature. *Frontiers in Nutrition, 6*(62). doi:10.3389/fnut.2019.00062

Vázquez-Fresno, R., Tulipani, S., Khymenets, O., Urpi-Sarda, M., Garcia-Aloy, M., Rabassa, M., . . . Andres-Lacueva, C. (2014). Emerging applications of metabolomics to polyphenols and CVD biomarker discovery. In R. R. Watson, V. R. Preedy, & S. Zibadi (Eds.), *Polyphenols in Human Health and Disease* (pp. 1025-1044). San Diego: Academic Press.

Vemireddy, L. R., Satyavathi, V. V., Siddiq, E. A., & Nagaraju, J. (2015). Review of methods for the detection and quantification of adulteration of rice: Basmati as a case study. *Journal of Food Science and Technology, 52*(6), 3187-3202. doi:10.1007/s13197-014-1579-0

Viola, R., Tucci, A., Timellini, G., & Fantazzini, P. (2006). NMR techniques: A non-destructive analysis to follow microstructural changes induced in ceramics. *Journal of the European Ceramic Society, 26*(15), 3343-3349. doi:https://doi.org/10.1016/j.jeurceramsoc.2005.09.055

Vogel, K. A., Martin, B. R., McCabe, L. D., Peacock, M., Warden, S. J., McCabe, G. P., & Weaver, C. M. (2017). The effect of dairy intake on bone mass and body composition in early pubertal girls and boys: a randomized controlled trial. *The American Journal of Clinical Nutrition, 105*(5), 1214-1229. doi:10.3945/ajcn.116.140418

Wang, Y., Liu, F., Li, P., He, C., Wang, R., Su, H., & Wan, J.-B. (2016). An improved pseudotargeted metabolomics approach using multiple ion monitoring with time-staggered ion lists based on ultra-high performance liquid chromatography/quadrupole time-of-flight mass spectrometry. *Analytica Chimica Acta, 927*, 82-88.

Ward, E. (2014). Addressing nutritional gaps with multivitamin and mineral supplements. *Nutrition Journal, 13*(1), 72. doi:10.1186/1475-2891-13-72

Wehrmueller, K., Jakob, E., & Ryffel, S. (2008). Orotic acid content in cowls, ewe's and goat's milk. *Agrarforschung, 15*, 356-360.

Wendorff, W. L., & Haenlein, G. F. (2017). Sheep milk – composition and nutrition. In *Handbook of Milk of Non-Bovine Mammals* (pp. 210-221).

Westerhuis, J. A., Hoefsloot, H. C. J., Smit, S., Vis, D. J., Smilde, A. K., van Velzen, E. J. J., . . . van Dorsten, F. A. (2008). Assessment of PLSDA cross validation. *Metabolomics, 4*(1), 81-89. doi:10.1007/s11306-007-0099-6

Whetstine, M. E. C., & Drake, M. A. (2006). Flavor Characteristics of Goat Milk and Other Minor Species Milk Products. In *Handbook of Milk of Non-Bovine Mammals* (pp. 107-120).

WHO. (2004). *Vitamin and mineral requirements in human nutrition*: World Health Organization.

Wiley, A. S. (2007). The globalization of cow's milk production and consumption: biocultural perspectives. *Ecology of Food and Nutrition, 46*(3-4), 281-312. doi:10.1080/03670240701407657

Wilson, B. (2008). The Swill is Gone. Retrieved from https://www.nytimes.com/2008/09/30/opinion/30wilson.html

Wishart, D. S. (2008). Metabolomics: applications to food science and nutrition research. *Trends in Food Science & Technology, 19*(9), 482-493. doi:https://doi.org/10.1016/j.tifs.2008.03.003

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*(2), 109-130. doi:https://doi.org/10.1016/S0169-7439(01)00155-1

Worley, B., & Powers, R. (2013). Multivariate analysis in metabolomics. *Current Metabolomics, 1*(1), 92-107. doi:http://dx.doi.org/10.2174/2213235X11301010092

Wu, D., Nie, P., Cuello, J., He, Y., Wang, Z., & Wu, H. (2011). Application of visible and near infrared spectroscopy for rapid and non-invasive quantification of common adulterants in Spirulina powder. *Journal of Food Engineering, 102*(3), 278-286. doi:https://doi.org/10.1016/j.jfoodeng.2010.09.002

Wu, J., Domellöf, M., Zivkovic, A. M., Larsson, G., Öhman, A., & Nording, M. L. (2016). NMR-based metabolite profiling of human milk: A pilot study of methods for investigating compositional changes during lactation. *Biochemical and Biophysical Research Communications, 469*(3), 626-632. doi:10.1016/j.bbrc.2015.11.114

Xi, X., Kwok, L.-Y., Wang, Y., Ma, C., Mi, Z., & Zhang, H. (2017). Ultra-performance liquid chromatography-quadrupole-time of flight mass spectrometry MSE-based untargeted milk metabolomics in dairy cows with subclinical or clinical mastitis. *Journal of Dairy Science, 100*(6), 4884-4896. doi:https://doi.org/10.3168/jds.2016-11939

Xin, H., & Stone, R. (2008). Chinese probe unmasks high-tech adulteration with melamine. *Science, 322*(5906), 1310-1311.

Xiu, C., & Klein, K. K. (2010). Melamine in milk products in China: Examining the factors that led to deliberate use of the contaminant. *Food Policy, 35*(5), 463-470. doi:https://doi.org/10.1016/j.foodpol.2010.05.001

Xu, J., Li, J., Zhang, R., He, J., Chen, Y., Bi, N., & Abliz, Z. (2019). Development of a metabolic pathway-based pseudo-targeted metabolomics method using liquid chromatography coupled with mass spectrometry. *Talanta, 192*, 160-168. doi:https://doi.org/10.1016/j.talanta.2018.09.021

Xu, X., Roman, J. M., Issaq, H. J., Keefer, L. K., Veenstra, T. D., & Ziegler, R. G. (2007). Quantitative measurement of endogenous estrogens and estrogen metabolites in human serum by liquid chromatography−tandem mass spectrometry. *Analytical Chemistry, 79*(20), 7813-7821. doi:10.1021/ac070494j

Yadav, S., Yadav, N. D. S., Gheware, A., Kulshreshtha, A., Sharma, P., & Singh, V. P. (2020). Oral feeding of cow milk containing A1 variant of β casein induces pulmonary inflammation in male balb/c mice. *Scientific Reports, 10*(1), 8053. doi:10.1038/s41598-020-64997-z

Yang, J., Zheng, N., Soyeurt, H., Yang, Y., & Wang, J. (2019). Detection of plant protein in adulterated milk using nontargeted nano-high-performance liquid chromatography–tandem mass spectroscopy combined with principal component analysis. *Food Science & Nutrition, 7*(1), 56-64.

Yang, Y., Zheng, N., Zhao, X., Zhang, Y., Han, R., Yang, J., & Zang, C. (2016). Metabolomic biomarkers identify differences in milk produced by Holstein cows and other minor dairy animals. *Journal of Proteomics, 136*, 174-182.

Yanibada, B., Boudra, H., Debrauwer, L., Martin, C., Morgavi, D. P., & Canlet, C. (2018). Evaluation of sample preparation methods for NMR-based metabolomics of cow milk. *Heliyon, 4*(10), e00856. doi:https://doi.org/10.1016/j.heliyon.2018.e00856

Zaalberg, R. M., Buitenhuis, A. J., Sundekilde, U. K., Poulsen, N. A., & Bovenhuis, H. (2020). Genetic analysis of orotic acid predicted with Fourier transform infrared milk spectra. *Journal of Dairy Science, 103*(4), 3334-3348. doi:https://doi.org/10.3168/jds.2018-16057

Zachar, P., Soltés, M., Kasarda, R., Novotny, J., Novikmecova, M., & Marcincakova, D. (2011). Analytical methods for the species identification of milk and milk products. *Mljekarstvo, 61*, 199-207.

Zhang, A., Sun, H., Wang, P., Han, Y., & Wang, X. (2012). Modern analytical techniques in metabolomics analysis. *Analyst, 137*(2), 293-300. doi:10.1039/C1AN15605E

Zhao, Y., Chen, H., Feng, J., Chen, Z., & Cai, S. (2017). 1H NMR-based compositional identification of different powdered infant formulas. *Food Chemistry, 230*, 164-173. doi:10.1016/j.foodchem.2017.03.020

Zhou, J., & Yin, Y. (2016). Strategies for large-scale targeted metabolomics quantification by liquid chromatography-mass spectrometry. *Analyst, 141*(23), 6362-6373. doi:10.1039/C6AN01753C

Zhu, D., Hayman, A., Kebede, B., Stewart, I., Chen, G., & Frew, R. (2019). 31P NMR-based phospholipid fingerprinting of powdered infant formula. *Journal of Agricultural and Food Chemistry, 67*(36), 10265-10272. doi:10.1021/acs.jafc.9b03902

Zhu, D., Kebede, B., Chen, G., McComb, K., & Frew, R. (2020a). Changes in milk metabolome during the lactation of dairy cows based on 1H NMR and UHPLC–QToF/MS. *International Dairy Journal, 111*, 104836. doi:https://doi.org/10.1016/j.idairyj.2020.104836

Zhu, D., Kebede, B., Chen, G., McComb, K., & Frew, R. (2020b). Effects of the vat pasteurization process and refrigerated storage on the bovine milk metabolome. *Journal of Dairy Science, 103*(3), 2077-2088. doi:10.3168/jds.2019-17512

Zhu, D., Kebede, B., Chen, G., McComb, K., & Frew, R. (2020c). Impact of freeze-drying and subsequent storage on milk metabolites based on 1H NMR and UHPLC-QToF/MS. *Food Control*, *116*, 107017. doi:https://doi.org/10.1016/j.foodcont.2019.107017

Appendix A

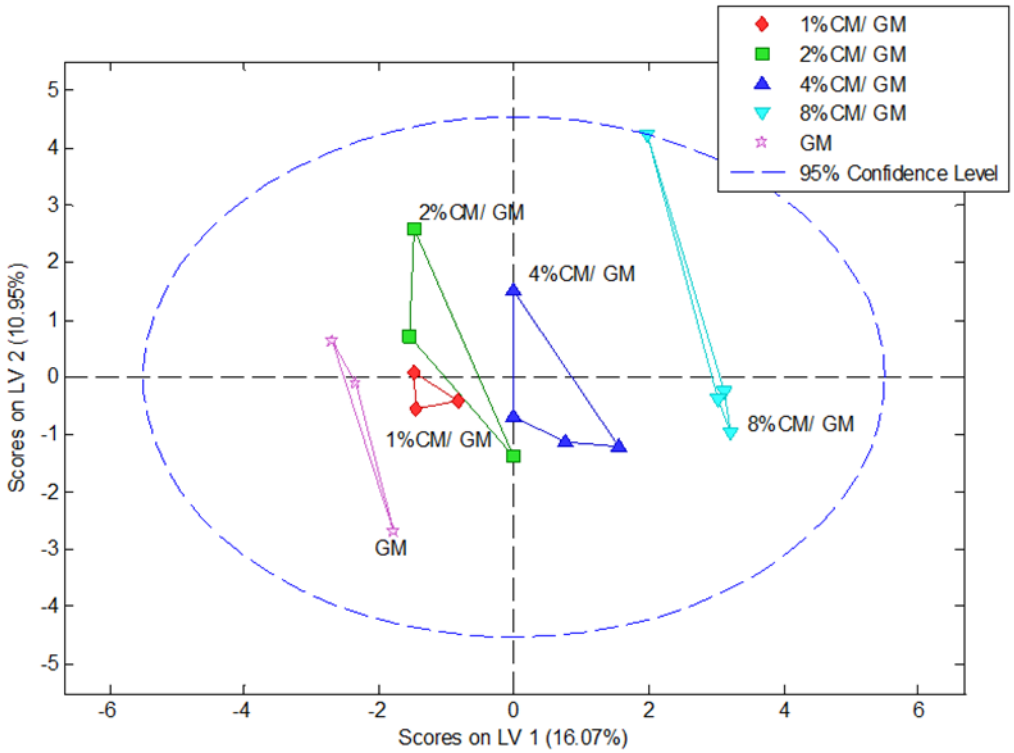

Figure A.1. PLS-R score plot for adulterated GM adulteration with different percentage of CM (1%,2%,4%, and 8%)

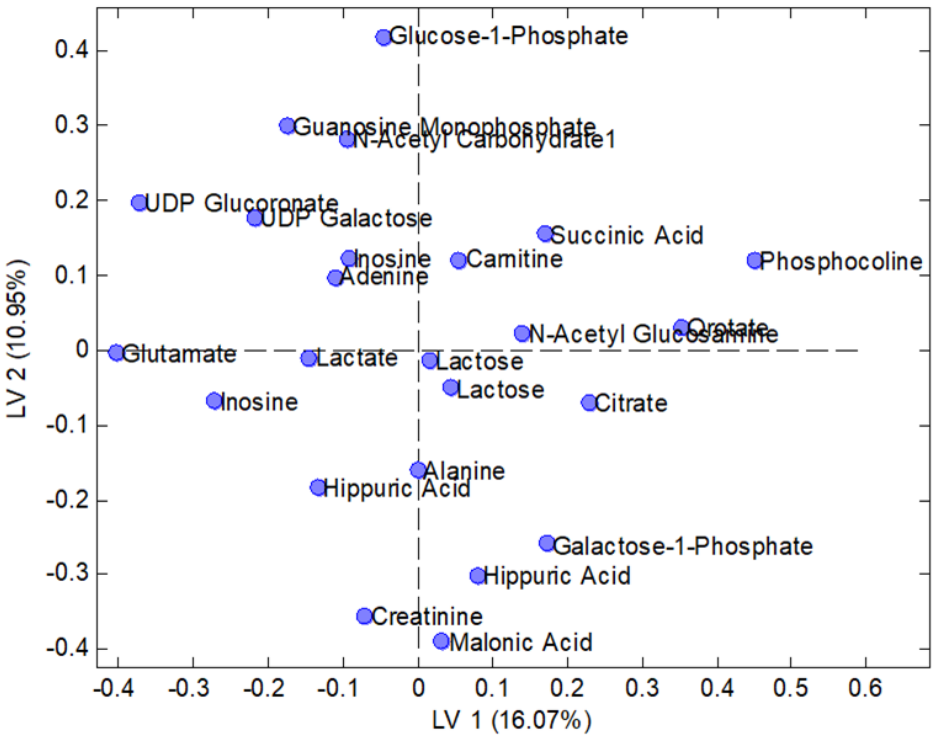

Figure A.2. PLS-R loadings plot for adulterated GM adulteration with different percentage of CM (1%,2%,4%, and 8%)

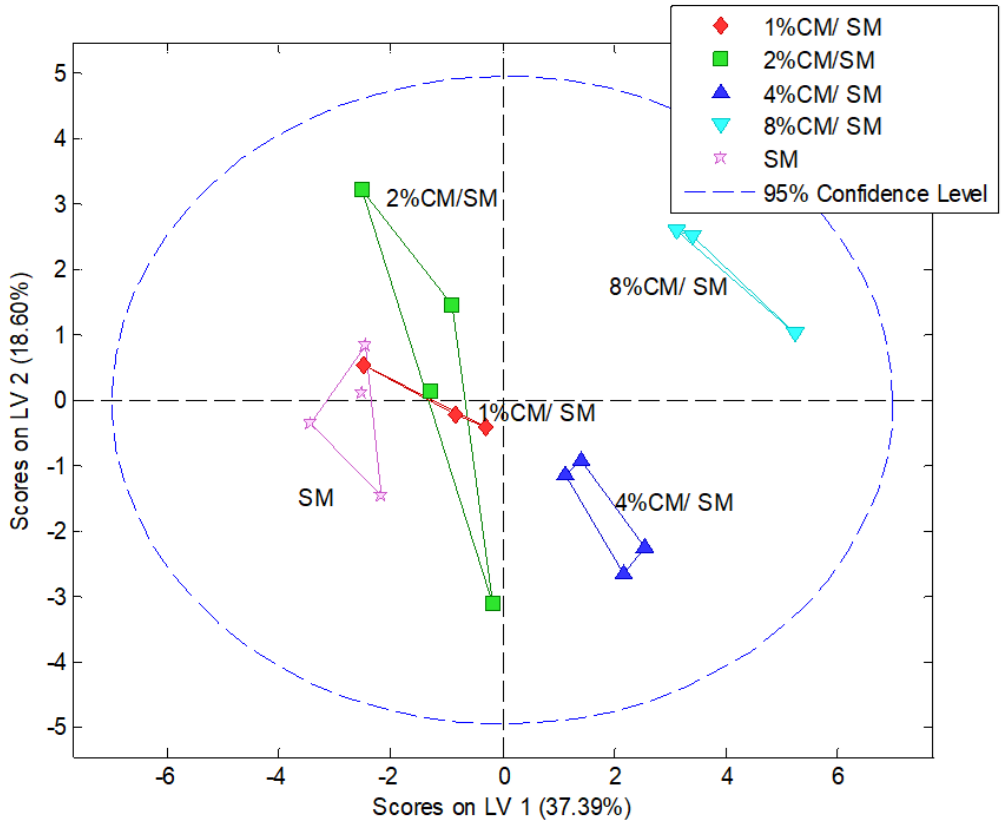

Figure A.3. PLS-R score plot for detecting SM adulteration with different percentage of CM (1%,2%,4%, and 8%)

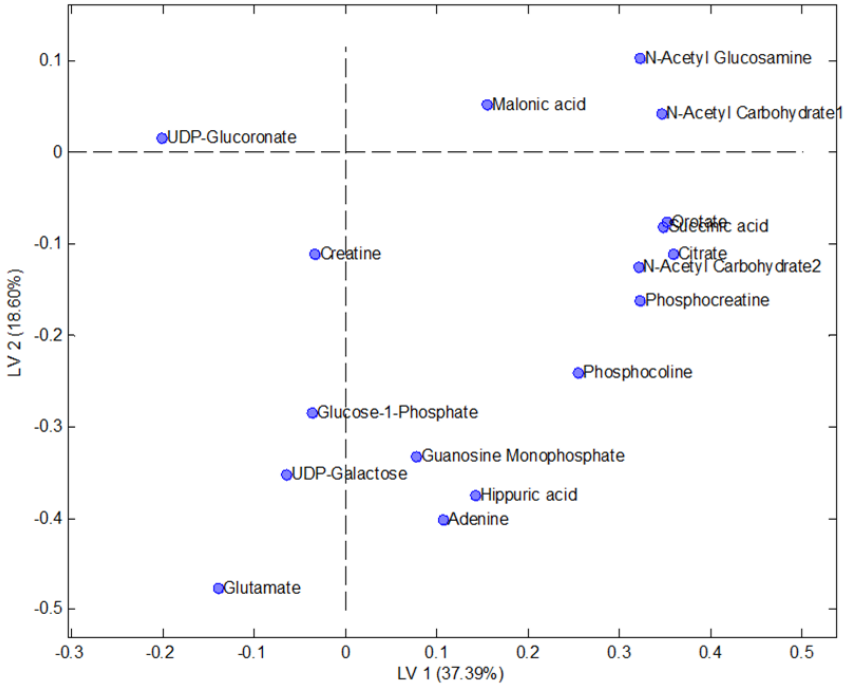

Figure A.4. PLS-R loadings plot for detecting SM adulteration with different percentage of CM (1%,2%,4%, and 8%)